Annual Report of Investigations Carried Out Under the Director's Research and Development Fund

September 28, 2009 to September 26, 2010

January 29, 2011

JPL
Jet Propulsion Laboratory
California Institute of Technology

FOREWORD

The Jet Propulsion Laboratory (JPL) Director's Research and Development Fund (DRDF) was established in 1969 as one provision of a Memorandum of Understanding between the National Aeronautics and Space Administration (NASA) and the California Institute of Technology (Caltech).

In 2010 the level of support for DRDF was $4.0 million. For a typical mix of in-house labor and procurements, this level of funding provides about 6.7 JPL work years per year of research effort plus the effort of our external collaborators. In contrast to the NASA civil service centers' flexibility of action afforded by the availability of research and program management (R&PM) resources, JPL must cover all cost of in-house labor, services, and related burdens by the R&D funds themselves.

The DRDF program has two components. Half of the program (normally called DRDF) is open to all domestic universities, companies, NASA Centers, and government laboratories. The other half of the program is called the Strategic University Research Partnership (SURP) program and concentrates on specific university partners, namely: Arizona State University, Carnegie Mellon University, Dartmouth College, Massachusetts Institute of Technology, Princeton University, Stanford University, University of Arizona, University of California at Los Angeles, University of Michigan, and University of Southern California. Both programs require external collaboration on research that offers a possibility of bypassing today's space mission limitations and can lead to revolutionary improvements in Earth and space science and missions. The broad DRDF component does not specify research topics but emphasizes innovative science and breakthrough technologies. The SURP component emphasizes specific strategic research topics plus an educational portion.

This annual report deals primarily with the technical content of the work supported by the DRDF. It is a collection of task reports prepared by individual investigators that describes the objectives, progress and accomplishments, significance of the results, and financial status of the various tasks. Care is taken by authors to credit other persons, especially university faculty and students, who contributed to the effort. Publications and conference presentations related to the work are listed at the end of each summary. The material has been lightly edited to ensure consistency in format, correctness of spelling and punctuation, completeness of thought, and so on, but the substance of style of the reports reflects the individuality of their authors.

The individual summaries are categorized as either "interim" or "final," according to the status of the task efforts. Most of the DRDF tasks that were active during FY 2010 are represented in the report, but in a few instances contributions were not forthcoming in time to be included in this volume. Submissions of this kind that eventually appear will be retained in the task files for reference.

The table of contents groups the summaries by research area.

Individual task reports are intended to be informative stand-alone narratives. However, readers who have questions or desire additional details about any task should not hesitate to contact the authors directly.

ACKNOWLEDGEMENT

CONTENTS

CONTENTS

INFORMATION SYSTEMS

MATERIALS

MICRO/NANO DEVICES

CONTENTS

CONTENTS

CONTENTS

PART 2: SURP ACTIVITIES

ASTRONOMY AND FUNDAMENTAL PHYSICS

COMPUTATIONAL SCIENCE

CONTENTS

CONTENTS

EDUCATION

CONTENTS

CONTENTS

GEOMETRY AND STRUCTURE OF OUR UNIVERSE

CONTENTS

CONTENTS

CONTENTS

PART 1: DRDF ACTIVITIES

ADVANCED MEASUREMENT TECHNIQUES

EFFICIENT ROOM-TEMPERATURE 3.0–3.5 mm SEMICONDUCTOR LASER FOR DETECTION OF METHANE AND HYDROCARBONS ON MARS AND TITAN

Alexander Soibel (PI), Instrument Electronics and Sensors Section (389)
Cliff Frez, Instrument Electronics and Sensors Section (389)
Siamak Forouhar, Instrument Electronics and Sensors Section (389)
Leon Shterengas, Department of Electrical Engineering, Stony Brook University
Gregory Belenky, Department of Electrical Engineering, Stony Brook University

A.   OBJECTIVES

The objective of this work is to develop high-performance semiconductor diode lasers operating in the 3.0–3.5 mm spectral range for planetary atmospheric studies and in particular for detection of hydrocarbons on Mars and Titan. Compared to the state-of-the-art interband cascade (IC) lasers delivered by us and incorporated in the tunable laser spectrometer (TLS) instrument scheduled for launch in 2009 on Mars Science Laboratory (MSL), the proposed lasers will have (a) 40–50 K higher operational temperature, and (b) more than 2 orders of magnitude better optical/electrical power conversion efficiency.

B.   APPROACH AND RESULTS

1.   Laser Design and Growth

We design and grow GaSb-based type-I QW diode lasers [1,2] employing quinary barrier material and operating in the spectral range near 3.15 μm. The laser active region comprised three 14-nm-wide, 40-nm-spaced, InGaAsSb QW with nominal indium composition of 52% and compressive strain of about 1.65%. The barrier and waveguide material was $Al_{0.20}Ga_{0.55}In_{0.25}As_{0.24}Sb_{0.76}$ quinary grown as a random solid alloy. Total width of the waveguide region (from n-cladding to p-cladding) including 3-QW active was about 1120 nm. The laser heterostructures were grown by solid-source molecular beam epitaxy using a VEECO GEN-930 reactor on tellurium-doped GaSb substrates. The n- and p-claddings were $Al_{0.85}Ga_{0.15}As_{0.07}Sb_{0.93}$ doped with tellurium and beryllium, respectively. Doping level of the part of the p-cladding layer adjacent to waveguide was kept around $2*10^{17}$ cm^{-3} to reduce internal losses associated with free hole absorption.

For initial material characterization, the part of the wafer was processed into multimode gain-guided lasers by opening 100-μm-wide windows in the dielectric followed by flat metallization. No etching was performed at this stage to facilitate either optical or electrical lateral confinement. Laser mirrors were cleaved. Uncoated 1-mm-long lasers were mounted epi-

side up for pulsed measurements. Figure 1 plots the light-current characteristics measured in the short pulse regime at different heatsink temperatures. The threshold current increases from below 200 mA to above 1 A when device operating temperature increases from 200 K to 290 K. The corresponding decrease of the laser slope efficiency is more than twofold. Strong temperature sensitivity of the laser's slope efficiency contributes to thermally activated increase of the laser threshold current. Poor thermal stability of the laser's external efficiency can originate either from internal optical loss increase or reduction of the injection efficiency with temperature. Unless free carrier absorption in active QWs is a dominant loss mechanism and/or QW carrier concentration is not pinned at threshold, neither of these effects are fundamental and, hence, temperature stability of laser threshold can be improved.

2.   Device Fabrication

Ridge waveguide lasers were fabricated using inductively coupled plasma (ICP) reactive ion etching technique. A very good control of the ridge sidewall was achieved during dry etching, as can be seen in the scanning electron microscope (SEM) image of the section of the ridge laser facet (Figure 2). The etching was stopped at 350–400 µm before the end of the p-cladding to form shallow 12-µm-wide ridge lasers with an estimated lateral effective index step of about 0.0002. For such a small refractive index step, the 12-µm-wide ridge waveguide is expected to support only fundamental mode. The ridge lasers were cleaved to form 2-mm-long cavities and coated to reflect about 30% (neutral reflection) and above 90% (high reflection), respectively. The distributed mirror losses of the coated 2-mm-long lasers are estimated to be about 3 cm^{-1}, i.e., lower then 12 cm^{-1} for 1-mm-long uncoated devices (Figure 1). Hence, the threshold carrier concentration is expected to be reduced in ridge lasers (as compared to test gain-guided devices) unless the ridge itself introduces extra losses.

3.   Laser Performance

Figure 3 plots the laser spectrum measured at different temperatures in pulsed mode (top) as well as current dependence of the CW spectrum at fixed heatsink temperature (bottom). The ridge laser spectrum got shifted to longer wavelength as compared to reference gain guided lasers. The Figure 3 data allow estimating the package thermal resistance as about 18 K/W. The values of the package thermal resistance we routinely obtain for 100-µm-wide 2-mm-long multimode lasers are about 2.5 K/W. Hence, one can conclude that thermal resistance for the devices with ridge width in the range from 10-µm to 100-µm roughly scales with ridge top area.

Figure 4 shows the light-current-voltage characteristics and slow axis far-field distribution measured at different temperatures. Lasers operate in the continuous wave (CW) regime up to 40 ºC with diffraction-limited output beam. The slow axis far-field distribution remains nearly unchanged in temperature range from 15 to 40 ºC. The measured FWHM is between 9 and 10 degrees. Model calculation predicted the FWHM about 8 degrees for a given ridge etch. The agreement between model and experiment is considered to be satisfactory since the model ignored the possible effect of the absorption loss outside the ridge structure and associated imaginary index step.

The maximum power of 9 mW near 3.18 µm was achieved at the current of 0.65 A at heatsink temperature of 15 ºC. The devices generate above 1 mW of CW power at heatsink temperature of 40 ºC (the operating wavelength is above 3.2 µm, see inset to Figure 4). The CW threshold current density was about 700 A/cm^2 at 15 ºC and it rising up to 2 kA/cm^2. The voltage drop across the laser heterostructure was below 2 V at the maximum output power level.

C.   SIGNIFICANCE OF RESULTS

These narrow-ridge waveguide index-guided mid-IR lasers exhibit much lower power consumption and can operate in a single spatial mode than is necessary for demonstration of single mode-DFB devices for spectroscopic applications. These lasers will enable a new generation of compact tunable diode laser spectrometers with lower power consumption, reduced complexity, and significantly reduced development costs compared to TLS on MSL. These lasers can be used for detection of HCN, C_2H_2, methane, and ethane. Measurements of these chemicals are of particular interest for Mars Exploration Program missions and future missions to Titan.

D.   NEW TECHNOLOGY

NTR number: 47377

E.   FINANCIAL STATUS

The total funding for this task was $200,000, of which $179,990 has been expended.

F.   ACKNOWLEDGEMENTS

G.   PUBLICATIONS AND PRESENTATIONS

[A]   J. Chen, T. Hosoda, G. Kipshidze, L. Shterengas, G. Belenky, A. Soibel, C. Frez, and S. Forouhar, "Single Spatial Mode Room Temperature Operated 3.15 µm Diode Lasers," *Electron. Lett.* **46** (2010): p. 367.

[B]   A. Soibel, C. Frez. S. Forouhar, A. Ksendzov, Y. Qiu, J. Chen, T. Hosoda, G. Kipshidze, L. Shterengas, and G. Belenky, "3.2 µm Single Spatial Mode Diode Lasers Operating at Room Temperature," accepted for oral presentation at CLEO 2010, San Jose McEnery Convention Center, May 16–21, 2010, San Jose, California

H. REFERENCES

[1]   L. Shterengas, G. Belenky, G. Kipshidze, and T. Hosoda, "Room Temperature Operated 3.1 µm Type-I GaSb-Based Diode Lasers with 80 mW Continuous-Wave Output Power," *Appl. Phys. Lett.* **92** (2008): p. 171111.

[2]   L. Shterengas, G. Belenky, T. Hosoda, G. Kipshidze, and S. Suchalkin, "Continuous Wave Operation of Diode Lasers at 3.36 microns at 12 degrees C," *Appl. Phys. Lett.* **93** (2008): p. 011103.

I. FIGURES
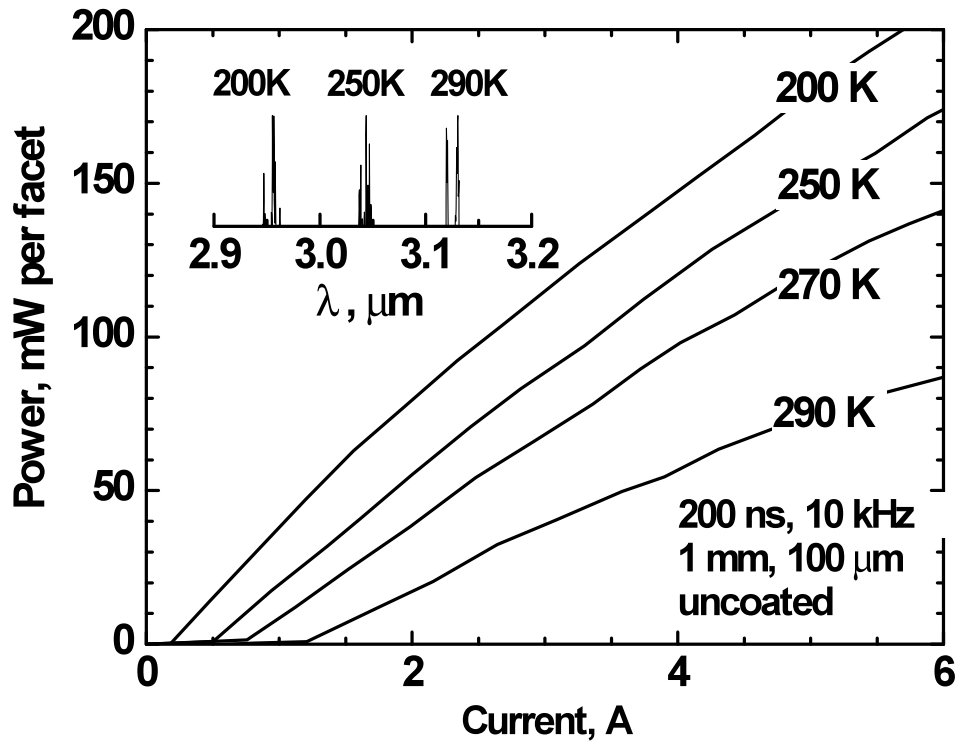
Figure 1. Temperature dependence of the short-pulse (200 ns, 10 kHz) light-current characteristics for 1-mm-long, 100-µm-wide, uncoated lasers. Inset shows the laser spectra.
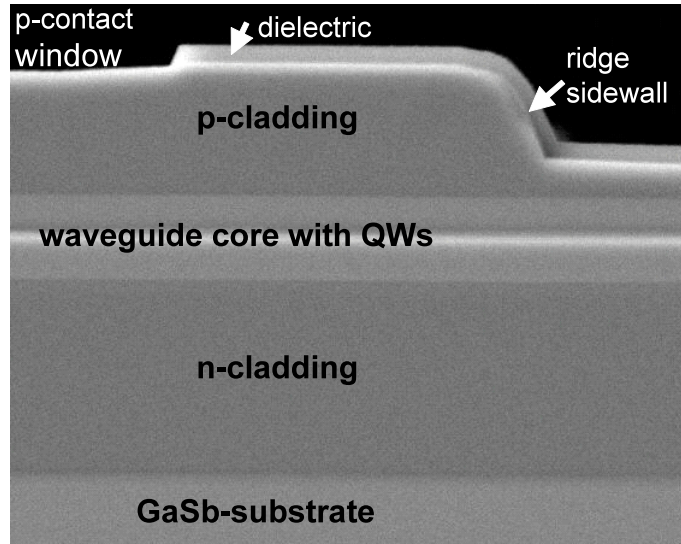
Figure 2. SEM image of the front facet section of the ridge waveguide laser.
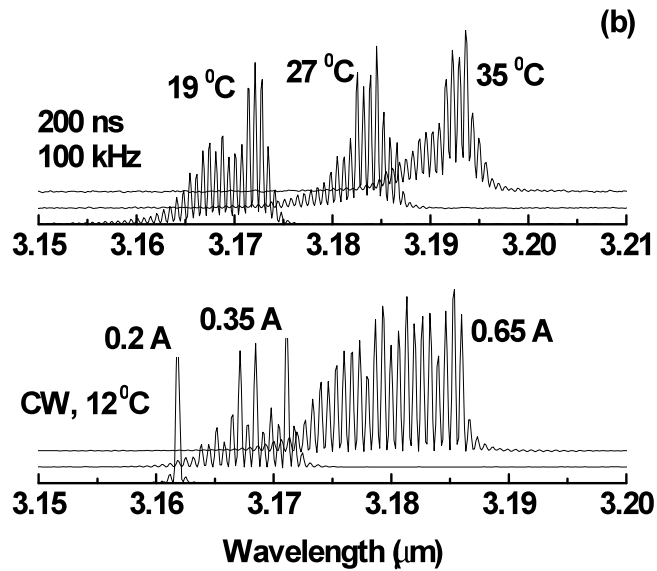
Figure 3. Ridge waveguide laser spectra measured in pulsed regime at different heatsink temperatures (top) and in CW regime at 12 ºC on heatsink. Spectra measured under different conditions are shifted vertically for clarity.
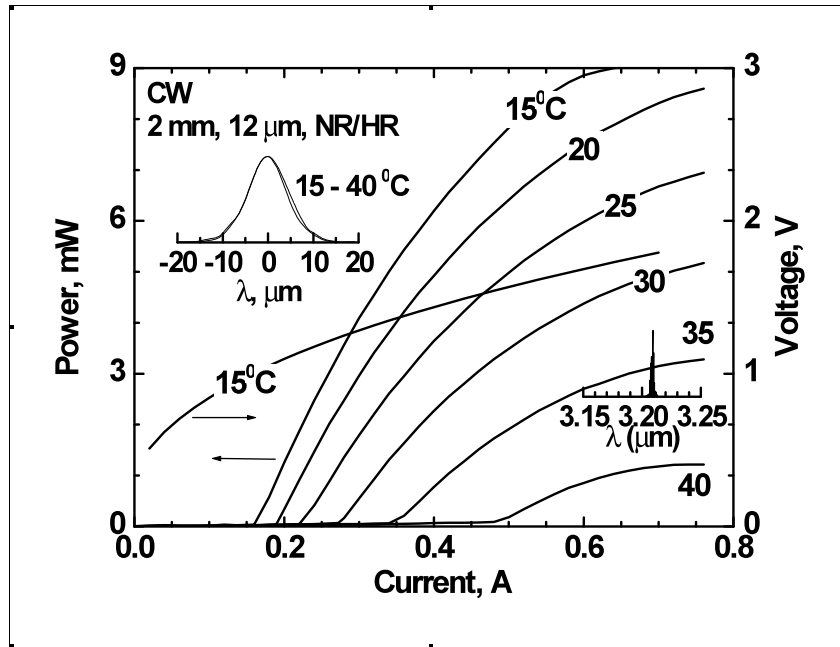
Figure 4. CW light-current-voltage characteristics measured for 12-µm-wide 2-mm-long ridge waveguide lasers in temperature range from 15 to 40 ºC. Inset shows the slow axis far-field distributions measured at 15 and 40 ºC. Laser spectrum at 0.4 A of CW current at 35 ºC at heatsink is also shown.

ADVANCED OPTICAL SYSTEMS

PRECISION PHOTOMETRY FOR DETECTION OF EARTHS ORBITING M STARS

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1343

Charles A. Beichman (PI), Astrophysics Science Centers (790)
GautamVasisht, Optical Section (383)
Roger Smith, Optical Observatories, California Institute of Technology

A.   OBJECTIVES

The goal of this activity was to develop a laboratory testbed to investigate the limits to photometric precision for near-infrared detectors. These detectors would be used for a variety of astrophysics programs, including detection and characterization of transiting planets, precision imaging of galaxies for weak lensing measurements, and precision radial velocity measurements of late type and/or young stars.

B.   APPROACH AND RESULTS

1.   Technical Approach

We used an existing Hawaii 2RG 2k × 2k detector available at Caltech to make a high-precision testbed. An existing dewar was modified to incorporate a stabilized LED, cryogenic integrating sphere, and a lenslet array to project over 14,000 small spots derived from a common illumination source (Figures 1 and 2). By monitoring the brightness of the spots as a function of time, we can now probe the limiting photometric performance at levels below 100 parts per million (ppm). Transit observations require performance levels of 10 to 50 ppm. By mounting the lenslet array on a flexure, we are able to shift the images of the spots across the detector array to simulate jitter due to pointing errors in a space telescope. By rotating the dewar, we have been able to control pointing offsets over a range of 15 pixels and with subpixel scale with a precision better than 0.01 pixel.

2.   Accomplishments to Date

We have recently completed the hardware system that is now taking data routinely in the laboratory. The dewar is very stable against external vibrations, which is critical for these high-precision observations. Drift is only 0.0016 pix//hr and stability over 5 minutes is 0.0001 pixel rms. We are presently analyzing the most recent data, which show that H2RG devices are capable of photometric precision in the 50 to 100 ppm range, or lower over a significant fraction of the detector (Figure 3).

C.   SIGNIFICANCE OF RESULTS

The demonstration of high-precision photometry with near-IR detectors will enable new space missions to detect transit or Earth-sized planets or to characterize them using spectroscopy. The results will be directly relevant for planning observations with the James Space Telescope (JWST) and eventually for weak lensing studies with the Wide Field Infrared Survey Telescope (WFIRST).

D.   FINANCIAL STATUS

The total funding for this task was $30,000 of which $23,364 has been expended.

E.   ACKNOWLEDGEMENTS

At Caltech, Gustavo Rahmer and Viswa Velur assisted with the development of the test dewar.

F.   PUBLICATIONS

This work is continuing with other funding sources and will lead to a publication in *PASP* later this year.

G.   REFERENCES

None.

H.   FIGURES

Figure 1. A schematic of the dewar as adapted for tests of precision photometry.

Figure 2. A subset of the 14,000 spots projected by the lenslet array.

Figure 3. Initial results from the testbed show photometric precision at 100 ppm.
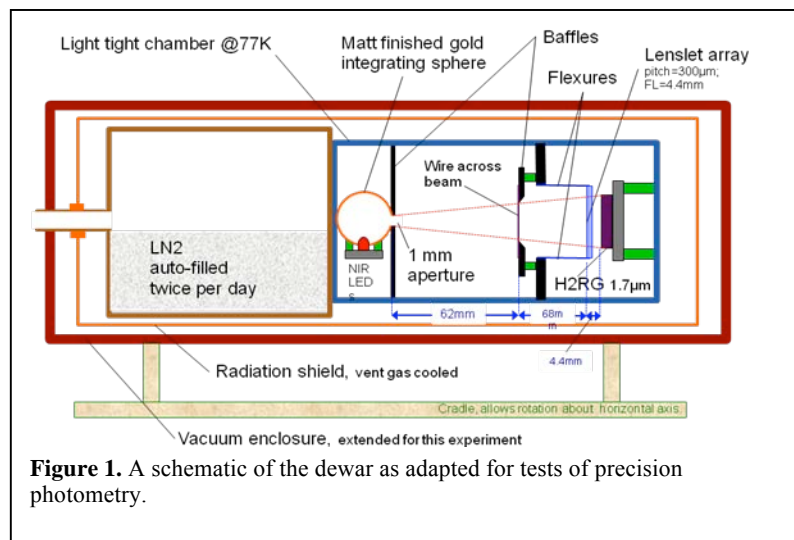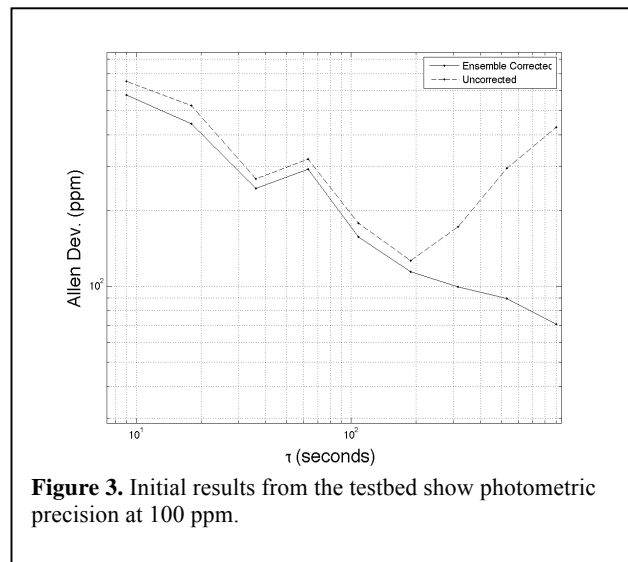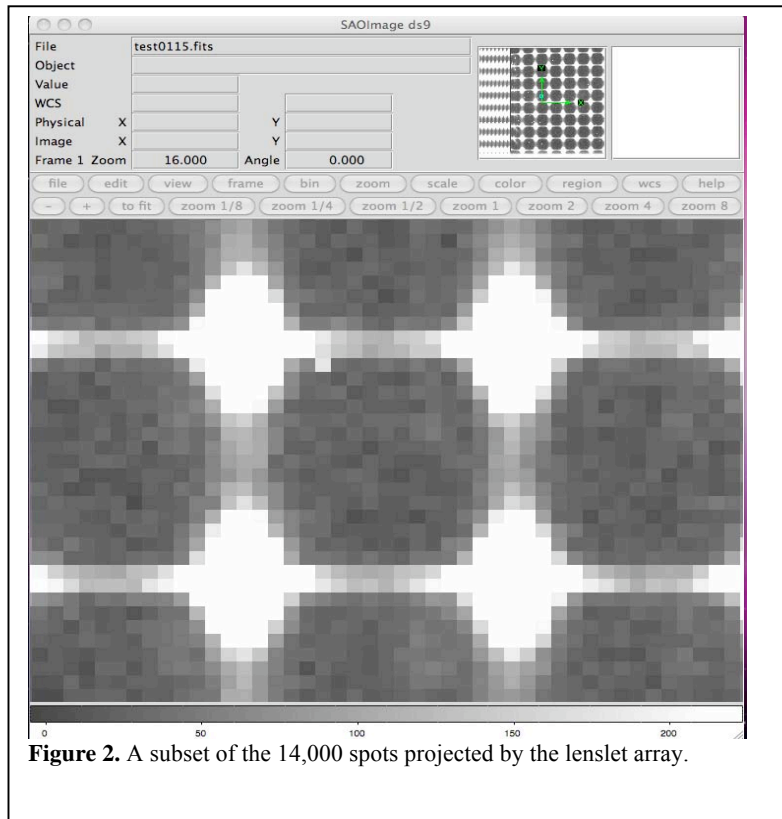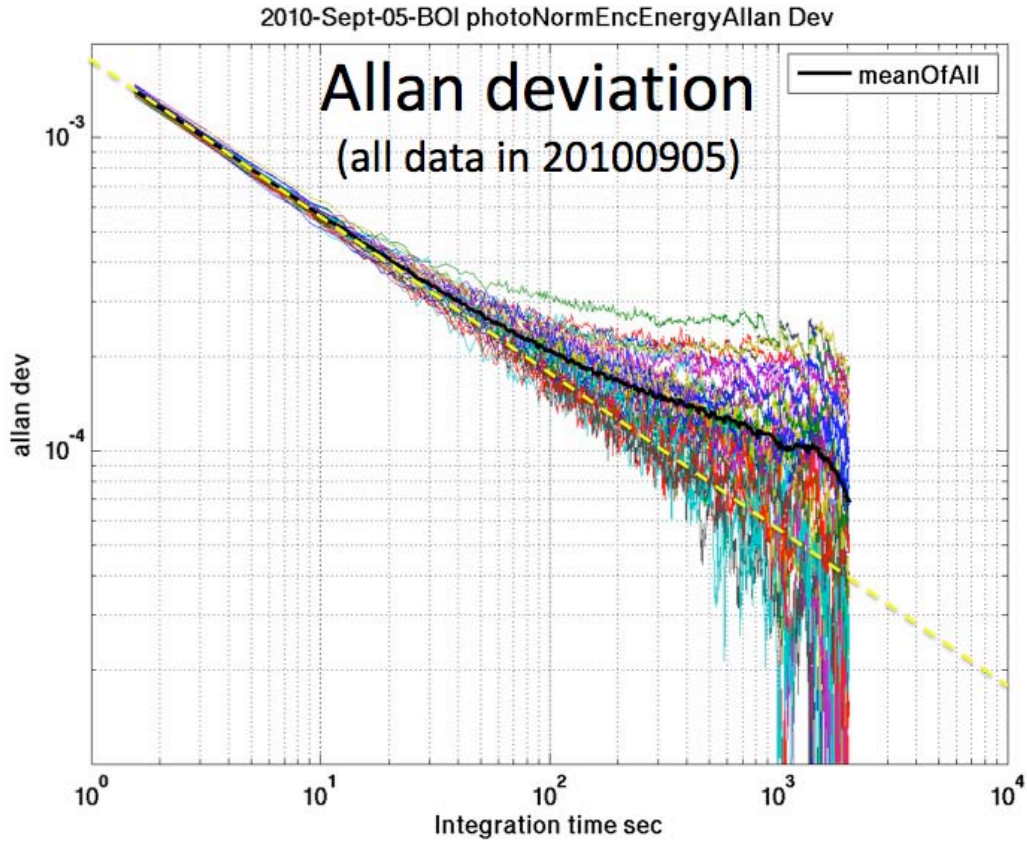
Figure 4. Allan Variance indicates the photometric precision as a function of total integration time after coadding. Different colors represent the scaling of photometric error for different spots, with the majority remaining close to the ideal shot noise down to 10^{-4} or lower. The mean behavior (black) is heavily biased by outliers, whose cause is under investigation and likely to be due to bad pixels, which have not been excluded here.

ASTROBIOLOGY

MINIATURE MANGANESE BIOMINERAL DETECTOR (MMBD) FOR MARTIAN EXOBIOLOGY

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1300

Soon Sam Kim (PI), Propulsion and Materials Engineering Section (353)
Susanne Douglas, Planetary Science (322)
Joseph L. Kirschvink, Division of Geological and Planetary Sciences,
California Institute of Technology

A.   OBJECTIVES

The objectives of this study are: (1) to develop a miniature manganese biomineral detector (MMBD) based on a wideband frequency scan electron paramagnetic resonance (EPR) spectrometer (2–6 GHz) capable of recognizing biomineral manganese oxides (Mn-oxides); (2) to probe the appearance of the Earth's oxidative atmosphere more than 2 billion years ago by analyzing ancient samples of Mn-oxides from strategic locations.

B.   APPROACH AND RESULTS

EPR spectroscopy is probably the most sensitive technique (for laboratory instruments @ $Q = 5000$, $\sim 10^{12}$ spins/cm^3 or 0.2 ng manganese oxides/cm^3) for studying atoms, molecules, or minerals with unpaired electrons, such as Mn-oxides [MnO_2, Mn(IV), 3d^3, S=3/2]. With application of magnetic field, the unpaired electrons split in energy (Zeeman interaction), and by microwave irradiation, one can induce transitions between the electronic spin states. To obtain EPR spectra, one can either scan the magnetic field (by an electromagnet requiring 5 kW power!) at a fixed microwave frequency (conventional EPR) or scan the microwave frequency at a fixed magnetic field. Our approach for the miniature EPR (5 W) is the latter, utilizing a frequency scan with a permanent magnet (Halbach magnet, @1300 Gauss, 350 g, with little fringe field).

Why Mn-oxide? It is believed that 4 to 3.5 byrs ago, when bacterial life first appeared on Earth, Mars had similar environmental conditions, including the presence of liquid water. Manganese is a relatively abundant element in the Martian regolith (0.45%), compared to 0.1% concentration in the Earth's crust. On Earth, the accumulation of manganese oxides in oceans, soils, sedimentary rocks, fresh water systems, and hydrothermal vents has been associated with microbial activity. Manganese is also a required trace nutrient for most life forms and participates in many critical enzymatic reactions such as photosynthesis. The widespread process of bacterial manganese cycling on Earth suggests that manganese is an important element to biology, and manganese biominerals represent potent biosignatures. Thus, it is reasonable to assume that life on Mars, if it ever existed, evolved in a similar fashion to that on Earth, and could have produced manganese biominerals that remain to this day as biosignatures. Such

manganese biominerals have a much better chance of survival in the Martian surface environment than do more conventional organic fossils. Iron oxides, such as hematite, are thought to be good indicators of the existence of past liquid water, but they are of much less value as biosignatures, having been formed abiologically at virtually any pH above 5. In marked contrast, manganese oxides are not at all common despite the fact that manganese is a relatively abundant metal on the Earth. In fact, compared to biogenic oxidation, abiotic homogeneous oxidation of Mn(II) to form Mn-oxides [MnO_2] is much slower up to pH values of 8.5, requiring years for completion at modest concentrations.

Until now, the structures of Mn-oxides have been obtained mainly through a suite of X-ray absorption spectroscopies (XAS), including X-ray absorption near-edge spectroscopy (XANES) and extended X-ray absorption fine structure (EXAFS). We have shown that EPR can contribute significantly to the characterization of Mn-oxides providing distinct spectral signatures for prescreening biominerals.

EPR Spectral Signatures. We have performed EPR studies on Mn-oxide samples from bacterial, synthetic analogs, biominerals, and mineral sources. Their EPR spectra are shown in Figure 1, plotted as the derivative of microwave absorption (at 9.3 GHz) vs. applied magnetic field, $\underline{H}_0$. The amount of samples used was typically in the range 0.5 to 3 mg. All the EPR data are plotted in Figure 2 as g values vs. spectral linewidths (ΔH). Notice the cluster of biogenic Mn-oxides around g = 2.0 and linewidths 400 Gauss < ΔH < 560 Gauss. By comparison, abiogenic Mn-oxides are widely scattered in g-values and linewidths, ΔH > 1200 Gauss. Suspected biomineral samples, Mn-nodules, desert varnish, and Kalahari show moderate linewidth, 600 Gauss < ΔH < 1200 Gauss. As far as we know, this is the first time that EPR has been applied in the search for Mn-oxide biosignatures. As noted in Figure 2, the linewidth criteria can be used to distinguish between abiogenic sources (blue) from the biogenic sources (green and red). For example, samples with broader linewidths (>1200 Gauss), including ones with high iron content, would be screened out. In fact, we can use EPR signatures as a convenient means of prescreening biogenic manganese minerals against abiogenic counterparts. Our EPR results match nicely with structural information obtained by X-ray techniques (see below). As such, it is an instrument of great value in search for biogenic manganese oxides on the surface of Mars.

Why the EPR Biosignature? Paramagnetic ions act as magnetic dipoles and the more they are in the proximity, the stronger they will interact, resulting in magnetic dipolar broadened spectra (ΔH_d). The importance of the dipolar contribution to observed EPR linewidth through interactions between Mn ions has been demonstrated in the case of alumina-supported manganese oxides (Mn / γ-Al_2O_3) used as catalysts, in which the Mn ions are dispersed on the surface of alumina [1]. These authors observed an EPR linewidth of 580 Gauss similar to the biogenic MnO_2 case at Mn loading of 0.5 wt%, with gradual broadening as the concentration of surface Mn ions was increased. The uniqueness of bacterial Mn-oxide resides in the fact that even though it is a neat Mn-oxide (100%), it can still present a narrow linewidth comparable to a very diluted counterpart.

All the Mn-oxide samples we have studied thus far show Lorentzian lineshape, which is characteristic of exchange-coupled paramagnetic moments with resulting linewidth, ΔH_e. The

observed EPR linewidths are results of such contributions, dipolar and exchange interactions. In general, interactions between identical ionic species, e.g., Mn(IV)-Mn(IV), result in exchange narrowing whereas interactions between different species, e.g., Mn(III)-Mn(IV), show exchange broadening [2]. Thus, Mn-oxide samples with more layers and mixed ionic states, i.e., Mn(III) and Mn(IV), will show a broader linewidth, ΔH, through exchange broadening, whereas samples with a small number of layers, holes in the layers, and little Mn(III) contents will display narrower linewidths through exchange narrowing. Thus, the widths of EPR spectra reflect the innate structure of Mn-oxides, including the presence and proximity of paramagnetic neighbors and presence of different ionic species, e.g., Mn(IV)-Mn(IV), Mn(III)-Mn(IV), which, through exchange interactions, either narrow or broaden the spectrum.

Key aspects of an abiogenic mineral birnessite structure include the presence of Mn(III) as a ubiquitous structural substituent in nominally defined as layered MnO_2, where it can occupy up to 1/3 of all nominal Mn(IV) sites [3]. At such high concentrations, Mn(III) is believed to order into rows parallel to the b crystallographic axis, which is accompanied by loss of the true hexagonal symmetry of the Mn oxide layer (and acquires a triclinic unit cell).

Bacteriogenic Mn-oxides produced by *Bacillus sp.* (SG-1), *Pseuodomonas putida* (GB-1), and *Leptothrix discophora* (SP-6) have recently been shown to be nanolayered Mn oxides, generally exhibiting a defected hexagonal birnessite structure [4,5,6] (Figure 1). In general, biogenic Mn oxides are believed to be relatively Mn(III)-poor (with average Mn oxidation state of 3.7–4.0) and Mn(IV) vacancies in fresh biogenic Mn oxides occur at apparent concentrations approaching 40 or 50% of all Mn sites [5]. These high vacancy concentrations are attributed in part to the nano-sized habit of the particles, which creates abundant layer terminations that appear as narrower linewidths in the EPR spectra. Mn(IV) site vacancies and edge sites are so abundant in biogenic Mn oxides, that the majority of occupied Mn(IV) sites are likely to reside in the immediate vicinity of a vacancy, whereas abiogenic mineral Mn oxides are likely to have the opposite character; a small fraction of occupied Mn(IV) sites are likely to reside next to vacancies. In principle, either of these structural aspects, i.e., high Mn(IV) layer vacancies or little Mn(III) content will account for the observed narrower EPR linewidth of the biogenic Mn-oxides. In a natural environment, once formed, the vacancies of the bacterial Mn-oxides will gradually be filled with various cations and may attain more layers with time. However, as Mn-nodules, desert varnish and Kalahari samples show, they still maintain narrower EPR linewidths than their abiogenic counterparts (Figure 2).

In summary, our EPR results are in good agreement with those from X-ray based structural characterizations. Thus, as long as biogenic Mn oxides consist of mostly single ionic species, e.g., Mn(IV), and lower numbers of layers with higher site vacancies, one can use corresponding EPR signatures as a basis for selection against abiogenic counterparts that contain more layers, fewer holes, and mixed ionic species.

Mn-Oxide in the Context of Planetary Evolution. A vital aspect of the evolution of life on Earth is the transition from anoxic to oxic life forms. Kirschvink and colleagues have predicted that major biologic mineralization of manganese began between 2.45 and 2.22 billion years ago, associated with rise in free oxygen in Earth's atmosphere. In this respect, biogenic Mn-oxides associated with late Precambrian (oxic) glaciations offer a standard for comparison with the

Paleoproterozoic (early Precambrian) manganese samples. Our trace element and new laser-ablation inductively coupled plasma mass spectrometry (ICP-MS) characterization of mass-independent oxygen isotope composition of Mn-oxide drill-core samples (never done before) reveal trace of free oxygen (molecular, O_2) associated with the least altered Paleoproterozoic glacial manganese formations from South Africa. So the preexisting hypothesis of biogenic contribution to these formations is now supported by geochemical data.

Results. We have developed the MMBD instrument based on a frequency-scan EPR spectrometer that can detect manganese biominerals. The EPR spectrometer utilizes a tunable loop-gap resonator (LGR, 2–6 GHz) that can cover the frequency ranges of Mn-oxides by mechanical sliding of a dielectric tuner over the resonator gap (Figure 3). The MMBD instrument produced frequency scan EPR spectra of DPPH (laboratory radical standard, g = 2.0037) and MnO as a proof of functionality (Figure 4).

C.   SIGNIFICANCE OF RESULTS

We have discovered that the EPR linewidths of bacteriogenic Mn-oxides are uniquely different from abiogenic Mn-oxides, being closely related to their innate crystal structure and Mn(III) content. This finding of distinct EPR spectral signatures of biogenic Mn-oxides is a unique JPL contribution to planetary exobiology, and breathes new life into the investigations using Mn-oxides as markers of modern and distantly past primitive biogenic activity. Even though it is widely believed that the Mn-oxides in modern natural waters and oxic surface environments are primarily the result of biological catalysis, until now there have been no criteria for confirmation of this belief. With the newly discovered EPR spectral criteria, we can shed light on the biogenicity of many suspected Mn-oxide minerals.

D.   NEW TECHNOLOGY

Under this DRDF, we have developed a wideband frequency scan (2–6 GHz) EPR spectrometer that can detect manganese biominerals. The spectrometer utilizes a tunable loop-gap resonator (LGR) by mechanical sliding of a dielectric tuner over the resonator gap (Figure 3).

E.   FINANCIAL STATUS

The total funding for this task was $195,000, all of which has been expended.

F.   ACKNOWLEDGEMENTS

PI Kim acknowledges contribution of JPL team member, Steven R. Carnes, for the development of wideband frequency scan LGR. Prof. Kirschvink acknowledges Dr. Timothy Raub for his efforts on characterization of geologic Mn-oxide samples.

G.   PUBLICATIONS

None.

H.   REFERENCES

[1]   W. S. Kijlstra, E. K. Poels, A. Bliek, B. M. Weckhuysen and R.A. Schoonheydt, "Characterization of Al_2O_3-Supported Manganese Oxides by Electron Spin Resonance and Diffuse Reflectance Spectroscopy," *J. Phys. Chem. B* **101** (1997): pp. 309–316.

[2]   W. S. Glaunsinger, H. S. Horowitz, J. M. Longo and A. Chang, "Magnetic Properties of the Mixed-Valence Compoounds $CaMn_3O_6$ and $CaMn_4O_8$," *J. Solid State Chemistry* **35** (1980): pp. 187–191.

[3]   B. Lanson, V. A. Drits, Q. Feng, and A. Manceau, "Structure of Synthetic Na-birnessite: Evidence for a Triclinic One-Layer Unit Cell," *American Mineralogist* **87** (2002): pp.1662–1671.

[4]   J. R. Bargar, B. M. Tebo, U. Bergamann, S. M. Webb, P. Glaetzel, V. Q. Chiu, and M. Villalobos, "Biotic and Abiotic Products of Mn(II) Oxidation by Spores of the Marine *Bacillus sp.* Strain SG-1," *American Mineralogist* **90** (2005): pp. 143–154.

[5]   S. M. Webb, B. M. Tebo, and J. R. Bargar, "Structural Characterization of Biogenic Manganese Oxides Produced in Sea Water by the Marine *Bacillus sp.*, Strain SG-1," *American Mineralogist* **90** (2005): pp. 1342–1357.

[6]   M. Villalobos, B. Lanson, A. Manceau, B. Toner, and G. Sposito, "Structural Model for the Biogenic Mn Oxide Produced by *Pseudomonas putida*," *American Mineralogist* **91** (2006): pp. 489–502.

[7]   A. Manceau and J. M. Combes, "Structure of Mn and Fe Oxides and Oxyhydroxides: A Topological Approach by EXAFS," *Phys. Chem. Minerals* **15** (1988): pp. 283–295.

I. FIGURES
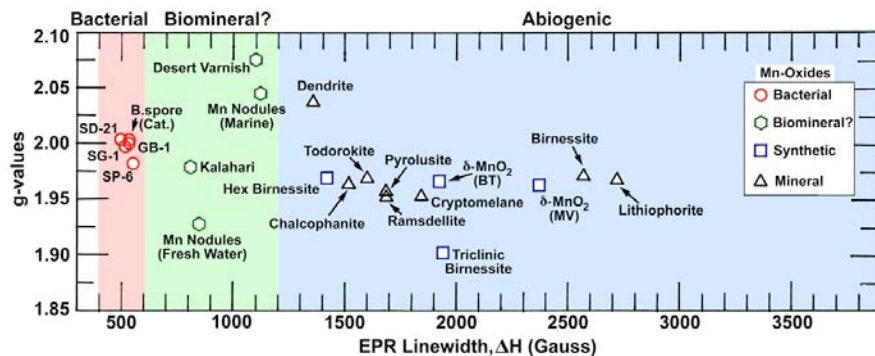
Figure 1. Left: Spores of the marine *Bacillus sp*. strain SG-1. Mn(II) oxidation occurs on the outermost layer of the spores. Middle: EPR of bacterial and mineral manganese oxide samples obtained by JPL laboratory spectrometer (Bruker Elexsys 500). The EPR spectrum of iron counterpart, Fe^{3+}, from a biogenic magnetite sample produced by GS-15 (biologically induced magnetite, g=2.12) is shown for a comparison. The red lines note for the g=2 position, and the two blue lines mark the corresponding frequency scan ranges (2-6 GHz, @1050 Gauss) of the miniature EPR spectrometer, MMBD. Right top: X-ray structure of hexagonal layer edge sharing [Mn(IV)-O_6] octahedra for synthetic δ-MnO_2 (top view) [6]. Right bottom: Layer structure of mineral birnessite (side view) [7].
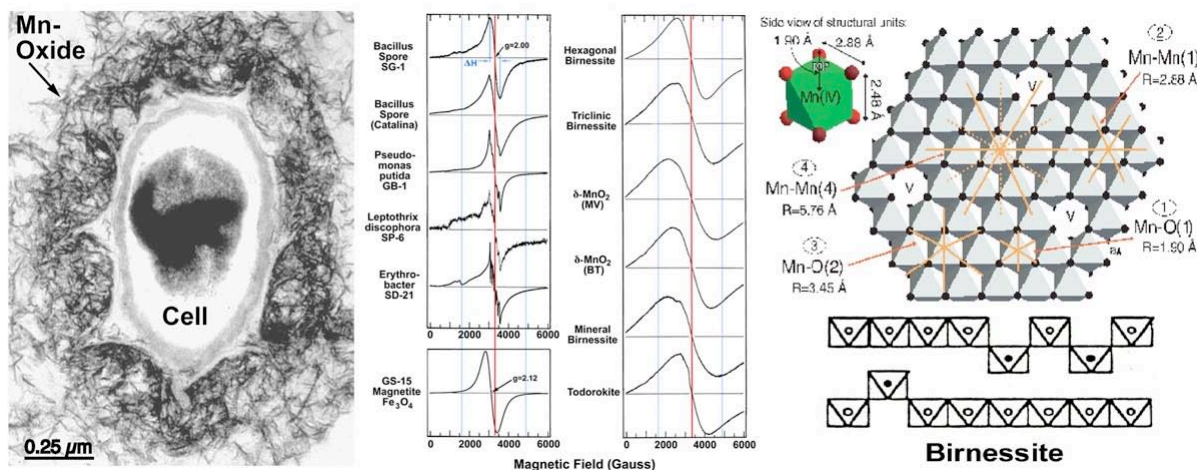
Figure 2. The EPR parameters of all the Mn-oxide samples studied are plotted as g-value vs. spectral linewidth. The bacterial ones show a narrow distribution in g-values and the linewidth, 400< ΔH < 560 Gauss, whereas the synthetic as well as mineral samples show a very wide distribution. Note the samples regarded as Mn-oxide biominerals; Mn-nodules, desert varnish, and Kalahari show narrower than the rest of the mineral samples.
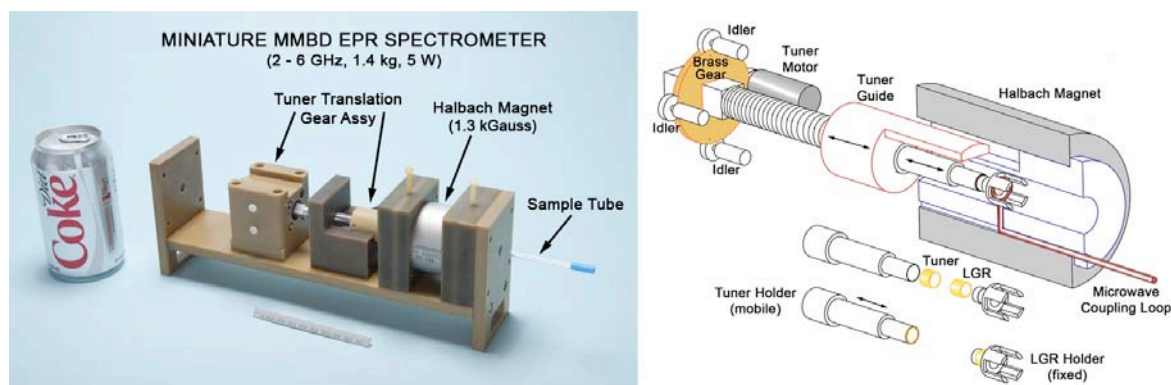
Figure 3. Left: Miniature MMBD wideband EPR spectrometer (2–6 GHz) developed under the DRDF. A soft drink can is shown for a size comparison. Right: A schematic drawing of the MMBD instrument. Details of the loop-gap resonator (LGR, 2–6 GHz) with the tuner (copper-backed dielectric Duroid, yellow) mounted on a tuner holder is shown. For clarity of presentation, a cutaway view is shown for the Halbach magnet.
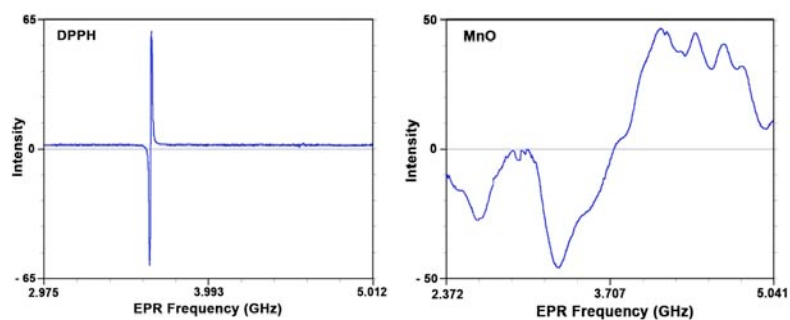
Figure 4. Frequency scan EPR spectra obtained by the MMBD instrument (Halbach magnet @ 1.3 kGauss). Left: DPPH (radical standard, g = 2.0037). Right: MnO.

ELECTRONICS

**RADIATION-HARD, DIGITAL CONTROLLED SINGLE-CHIP ALTERNATIVE
TO HYBRID SWITCHING REGULATORS**

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1302

Philippe C. Adell (PI), Electronic Parts Engineering Office (514)
Bert Vermeire, Electrical Engineering, Arizona State University
Bertan Bakkaloglu, Electrical Engineering, Arizona State University
K. K. Ma, Sandia National Laboratory

A.   OBJECTIVES

The objective of this work is to demonstrate the feasibility of a cost-effective, highly reliable, radiation-hardened, single-chip DC-DC switching power converter compliant to JPL mission requirements. We designed, fabricated, and tested an innovative digitally controlled converter for point of load (POL) application with built-in self-test (BIST) diagnostics and load characterization using a commercial foundry.

The goals were:

•   Design, fabrication, and testing of a digital controlled single-chip buck converter prototype using a non-hardened foundry. This collaborative research effort between JPL and ASU resulted in one of the most aggressive topologies ever reported.
•   Study the feasibility of transitioning the prototype to the $1B MESA fabrication facility at Sandia National Laboratory (SNL), a strategic JPL alliance.

B.   APPROACH AND RESULTS

1.   Approach

a.   Design of the scalable DC-DC converter module

A basic representation of the converter module with direct main power connectivity, self-diagnostics and adaptive digital control is shown in Figure 1. The digital control system operation is as follows: after voltage-to-frequency conversion using ring oscillator based VCOs, single-bit sigma-delta ($\Sigma\Delta$) modulated feedback signals are compared to the $\Sigma\Delta$ modulated analog voltage reference. The 3-level first order $\Sigma\Delta$ noise shaped error signal is then decimated using 2-stage comb (CIC) filter and applied to the compensator (PID) input. The PID compensator calculates the required duty cycle. Finally, the digital pulse-width modulator (DPWM) converts this voltage command into a duty cycle to drive the PFET and NFET via a built in non-overlap dead-time gate driver. The implementation of the converter architecture is based on the following modules:

• Digitally controlled DC-DC Converter: a digital PWM DC-DC buck converter utilizing a first-order $\Sigma\Delta$–based frequency discriminator (SDFD) is the core converter for this integrated circuit. A frequency discriminator generates an accurate representation of an instantaneous frequency of a carrier signal.

• Digital PWM Generators: We use a coarse and fine scheme to generate the high-accuracy PWM duty cycle. The coarse scheme use a counter-based approach, while the fine control is through a high-accuracy phase locked digitally controlled ring oscillator, i.e., DLL. This DLL controls the 4 fine bits of DPWM codes. Figure 2a shows the overview of the implementation.

• Digital Self-Diagnostic Circuits and Optional Digital Current Sensing Techniques: As shown in Figure 2(b), a self-diagnostic circuit is used to measure the Resr and L values of power inductor during converter operation. We built an on-chip triangular current generator, and pushed it into the inductor while shorting the power capacitor. An OPA, VCO, and Σ-Δ based frequency discriminator (SDFD) chain is used to read out and digitize the voltage over the inductor. The readout chain is similar with the VCO SDFD chain in digital control loop shown in Figure 1.

b. Transition to a non-hardened foundry: assessment

To match JPL mission requirements (i.e., Europa) and because power converters are very sensitive to total ionizing dose and single event effects, we studied the feasibility of transitioning the current converter prototype to the Sandia National Laboratory (SNL) rad-hard foundry. This evaluation involved analysis of current test data and test structure for process monitors for radiation performance.

2. Results

Two generation designs were developed during the course of this project. The first generation was fabricated in a commercial foundry and had lower input voltages and lower current capabilities, but validated experimentally the DPWM generation and digitally intensive feedback digitizers. The second generation implemented the digital self-diagnostic circuits, had higher current capabilities, and had higher voltage ratings. To minimize risk, the PID block for each generation was implemented through the use of an FPGA so that it provides design flexibility, particularly for adjusting design parameters to control loop gain and phase converter frequency response. PID implementation can be easily added to the chip through RTL synthesis and digital cell libraries.

• *First-generation design*: The DPWM generation and digitally intensive feedback digitizers have successfully been demonstrated experimentally. The converter results are shown in Figures 3 and 4. Figure 3 shows a closed-loop measurement; the converter output was recorded and regulated properly. The output ripple value was well below 1%. Figure 4 shows the transient output characteristics of the converter, with a full step load change, the converter output recovered within 150 µs and its transient change stays within 5% of the output value. The design shows maximum peak efficiency of 92% and through the full load.

• *Second-generation design:* Figures 5–8 give results about the second-generation design with higher voltage ratings, high load current characteristics, and digital self-

diagnostic circuits. The design has been fabricated (Figure 5), and successfully simulated (Figures 7 and 8). Electrical characterization with the PCB test board developed (Figure 6) is in progress at this time. Results will be released April–May timeframe and presented at the upcoming research and technology development conference.

      •   *Sandia process evaluation:* Sandia offers a unique and attractive 1 Mrad process for the current approach. It integrates 0.35 μm silicon-on-insulator (SOI) transistors with high-voltage drivers offering low specific-on-resistance N-channel LDMOS devices at up to 30 V. Process characteristics and radiation data are summarized below:

      a.   <u>LDMOS device</u>: A very attractive lateral DMOS device with a breakdown voltage of 30 V, 10 A and a good gate charge-on-resistance figure of merit (150 nC.mΩ). FY'09 development led to a device that is immune to single-event latchup (SEL), burnout (SEB), and gate rupture (SEGR); with a threshold LET (LET_{th}) > 80 $MeV.cm^2/mg$. These improvements impacted total hardness; however, as shown in Figure 9, the device shows little degradation up to 100 krad and the threshold voltage is still positive after 500 krad up to 1 Mrad.

      b.   <u>CMOS7</u>: A rad-hard, mixed-signal CMOS SOI (BUSFET) technology, 0.35 μm and 3.3 V operations. As shown in Figure 10: 1) total dose up to 500 krad with no parametric degradation and > 1 Mrad with parametric degradation, no functional failure; 2) Single-event effects (SEE) SEU threshold LET ~30 $MeV.cm^2/mg$.

      c.   <u>Available IP blocks</u>: Several IP blocks were evaluated against total dose. These blocks can be used for the final Sandia prototype, particularly voltage reference as input to VCOs and PLL/oscillator for digital clock generation. The IP blocks available are compatible with the current process design kit; allowing design development while using existing IP blocks. Table 1 summarizes the results.

C.   SIGNIFICANCE OF RESULTS

The industrial supply base for space-grade DC-DC converters has had routine difficulty meeting the reliability, schedule, and manufacturing standards set forth by the space industry. This effort develops a digitally controlled single-chip solution, thereby eliminating the principal production problems associated with the customary hybrid construction. The targeted 300 krad(Si) (goal 1 Mrad(Si)) power converter, with its improved reliability and power efficiency, will be industry leading. Although all missions will benefit from the reduced cost and schedule risk typically associated with commercial power converters, this design will be crucial for the planned Europa mission. Its scalable digital control approach lends itself to future applications that will require higher bus voltage (i.e., 70 V). We are targeting a design solution that can be used as an alternative to hybrid COTS converters in JPL power distribution units with the advantages of design flexibility, with similar or even better performance, reliability, and radiation-hardened. High-efficiency POL converters will greatly improve the overall efficiency of power systems, thereby significantly reducing power requirements for all JPL missions.

D.   FINANCIAL STATUS

The total funding for this task was $194,100, all of which all has been expended.

E. ACKNOWLEDGEMENTS

F. PUBLICATIONS AND PRESENTATIONS

[A]   Bertan Bakkaloglu, "Digital Controlled Buck Converter Using Frequency Domain ΣΔ ADC and Hybrid PWM Generator," presented at the International Solid-State Circuits Conference, San Francisco, California, February 7–11, 2010.

G. FIGURES

Table 1. TID response of Sandia IP blocks related to the converter design.

Design Blocks	TID level	Comments
1.2 V bandgap reference	5% degradation at 1 Mrad	Block will be used for final design and feed VCOs
PLL/oscillator	@1 Mrad less variation than temperature	Could be used as frequency generator for digital clock
ADC	@1 Mrad slight change in offset and gain	Transfer characteristics quite linear and functional
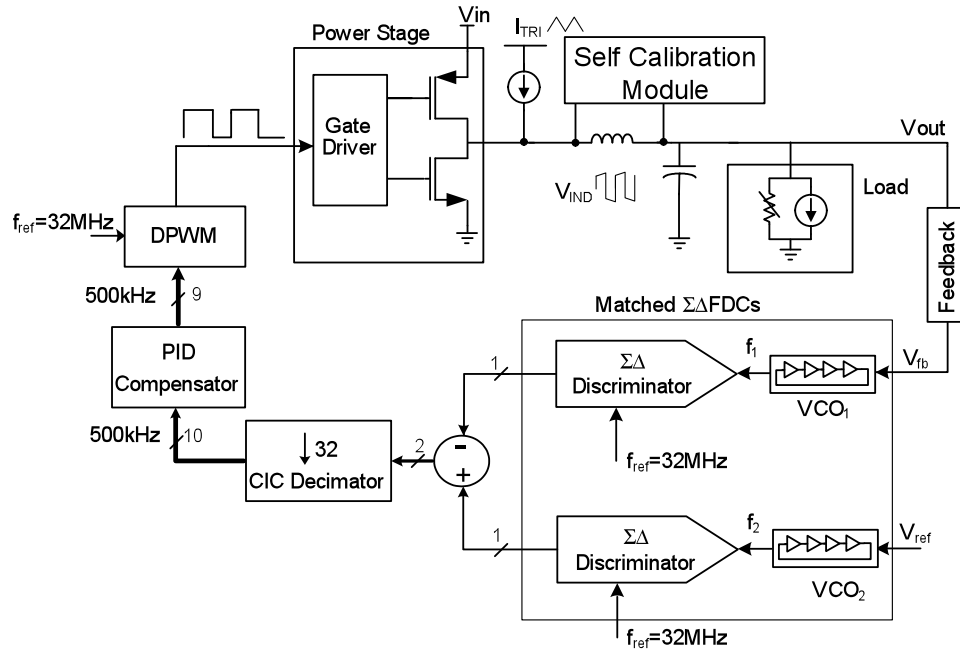
Figure 1. Proposed digitally controlled DC-DC converter with self-calibration module.

(a)

(b)

Figure 2. (a) Block diagram of digitally controlled PWM generator utilizing a coarse–fine duty cycle generation scheme; (b) built-in load diagnostics testing circuits.

Figure 3. Experimental validation of the closed loop design: output voltage ripple, and inductor current at full load condition. The ripple is below 1%.

Figure 4. Experimental validation of the closed loop design: 1) Transient response to a 300-mA load step; 2) voltage drop is around 130 mV from the regulated level of 1.8 V; and 3) control loop recovers with 150 μs settling time.
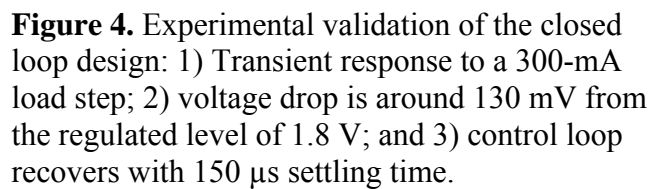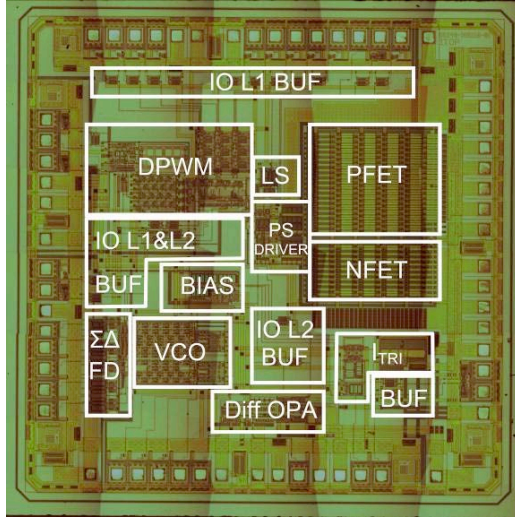
Figure 5. Test chip micrograph of the second-generation design, integrating higher current FETs and BIST. The AMIS i2t100 0.7 µm 3M process was selected for high-voltage, large-current requirements. The chip has 3.5×3.5 mm^2 and 61 pins.
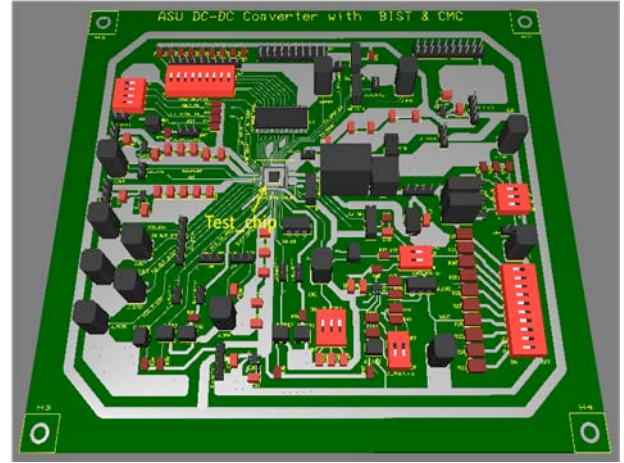
Figure 6. 3D representation of the PCB test board. With this board, comprehensive tests of the chip can be performed. It includes DC-DC normal operation test, load current sensing test, BIST mode inductor inductance test, efficiency test, and radiation test.
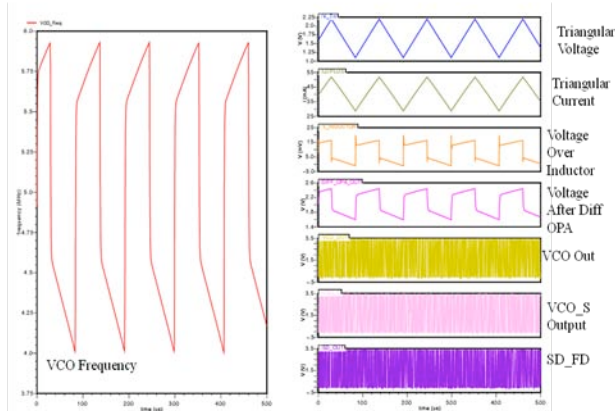
Figure 7. BIST for inductor simulation. A high linear triangular voltage was generated first and then converted to a triangular current though voltage to current converter. An OPA, VCO, and Σ–Δ based frequency discriminator (SDFD) chain is used to read out and digitize the voltage over the inductor.
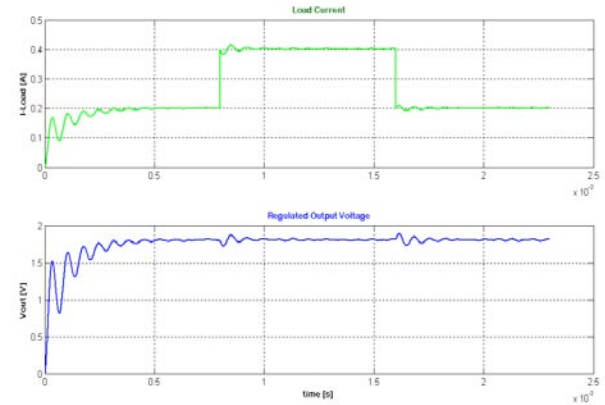
Figure 8. Closed-loop transient simulation with load step from 200 mA to 400 mA was shown in Figure 8. DC-DC regulator settling voltage target was set to 1.8V. It settles to desired voltage around 200 µs.
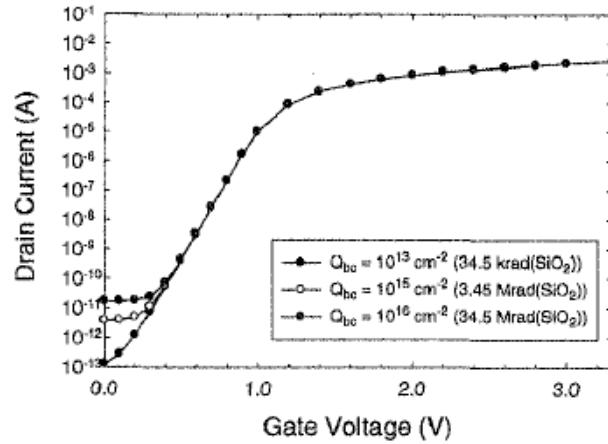
Figure 9. The simulated response of a partially-depleted SOI BUSFET to radiation-induced charge buildup at the back-channel interface. Subthreshold I-V characteristics as a function of back-channel interface charge density. Devices are radiation hard up to 1 Mrad.
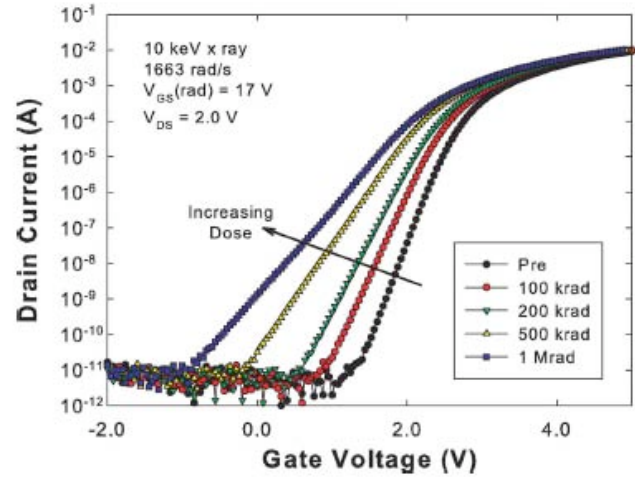
Figure 10. Wafer-level X-ray total dose response of lateral power MOSFET transistors. Characteristics: up to 500 krad, the threshold voltage remains positive, a critical parameter to ensure converter operation.

DEVELOPMENT OF THE HIGHLY EFFICIENT RAD-HARD-BY-DESIGN MICROPROCESSOR FOR ENABLING SPACECRAFT (HERMES)

Director's Research and Development Fund (DRDF)
Interim Report

JPL Task #1354

Steven M. Guertin (PI), Electronic Parts Engineering Office (514)
Michael A. Newell (Co-PI), Flight System Avionics Section (345)
Philippe C. Adell (Co-PI), Electronic Parts Engineering Office (514)
Raphael R. Some (Co-PI), Flight System Avionics Section (345)
Lawrence T. Clark (Co-PI), Department of Electrical Engineering, Arizona State University
Dan W. Patterson (Co-PI), Department of Electrical Engineering, Arizona State University

A.   OBJECTIVES

The Highly Efficient Radiation-Hardened by Design (RHBD) Microprocessor for Enabling Spacecraft (HERMES) is being studied for potential JPL uses and to examine the possibility of improving it to meet mission requirements, including verification of the technology at low temperature, operation at low power, and the addition of a floating point unit (FPU) to support applications. We are performing validation tests and trade studies of HERMES for this project.

HERMES is being developed at Arizona State University (ASU) under contract with the Air Force Research Laboratory (AFRL) and provides an opportunity for JPL to participate in development of appropriate RHBD methods and usage scenarios for a very low-power device.

The research areas pursued in this DRDF are the following:

•   HERMES is based on high-quality novel RHBD methods that will provide devices with higher single-event effects (SEE) hardness than current similar devices. Validation of the SEE sensitivity of test chips to proton radiation and examination of mechanisms relevant to SEE in deep submicron devices was performed using proton exposure.

•   An FPU will be required for any JPL application. We are examining methods for FPU implementation and comparing possible hardware and software implementations. As part of this work, the HERMES microarchitecture is being extended to support an FPU architecture to be chosen and implemented at a later time.

•   As part of a trade study, findings indicated that low-temperature validation of HERMES is of high interest to potential JPL users (e.g., applications in shaded lunar craters). Environmental verification of HERMES test chips will be performed at very low temperature (40 K).

• A trade study of implementation architectures for computers built using HERMES is underway. A HERMES-based computer for use at JPL will likely target very low-power applications. Two generic architectures are being studied for power and performance capabilities.

B. APPROACH AND RESULTS

1. Proton Testing of HERMES Test Chips

There are two HERMES test chips, TC19 and TC20, fabricated on the IBM 90 nm LP and SF processes, respectively. Both contain an RHBD high-performance 16 kB, 4-way set associative cache and a programmable built-in-self-test (BIST) engine to stimulate and measure the response of the cache at speed on the tested die [1]. The BIST engine is fully triple modular redundant (TMR) through the I/Os, so that the test engine does not hinder measurements of the circuits under test [2,3]. However, the PLL is unhardened IP from IBM. TC20 also has a prototype dual redundant register file (RF) and ALU/bypass loop, where the dual pipelines are checked against each other at the write-back state in the RF. An extra read port is used to save the data that is to be replaced on a write-back, for the case that an error is detected in the write operation itself; in this case, the machine architectural state is restored and the operation is restarted from the beginning of the pipeline. All circuits are designed using annular gate topologies and the test chips have been tested to greater than 2 Mrad(Si) TID with only negligible standby current increase.

In FY'10, proton testing of two HERMES test chips, TC19 and TC20, was performed. The test board is shown in the LBL beam line in Figure 1. This section reports preliminary results from proton testing at LBL in October 2010. Beam energies of 49.3 and 13.5 MeV were used. Most of the testing was performed at 60 MHz using the PLL bypass mode, since PLL upsets can cause major test chip malfunctions due to bad clocks. The bypass is non-TMR, so bad clocks can still occur, and some results seemed to indicate that data corrupted in both dual redundant copies, which seemed to indicate truncated clock cycles, i.e., the data transitions appear not to have finished.

For most RF tests, the total fluence was $5(10^{11})$ protons. For the 49.3 MeV protons, the measured cross-sections for static single event upset (SEU) testing of the RF was $3.71(10^{-14})$ cm^2, $2.68(10^{-14})$ cm^2, and $1.12(10^{-14})$ cm^2 at $V_{DD} = 1$, 1.2, and 1.4 V, respectively. Static tests write a pattern to the RF, and the BIST engine dumps the RF contents continuously. In the event of an error, it is corrected and the test continues at the next location. An SET, presumably in the RF dynamic read circuits, is detected if the read data is incorrect, but is correct on a subsequent read. Similar response was measured for all 0 and all 1's patterns. This was expected, since the RF was designed with essentially balanced capacitance on the two sides of the storage nodes. SETs were observed only for the read all 1's case but only in one test and with a very low cross-section. For the 13.5 MeV protons, the measured cross-sections for static SEU testing of the RF was $1.07(10^{-14})$ cm^2, $8.98(10^{-15})$ cm^2, and $1.72(10^{-15})$ cm^2 at $V_{DD} = 1$, 1.2, and 1.4 V, respectively.

The FPU pipeline (see Figure 2) runs independent of the integer pipeline; e.g., FPU operations continue during integer stalls. The RF/ALU bypass loop allows testing of the RF and dual redundant ALU under realistic microprocessor operating conditions; i.e., with instructions running and data flowing through the DUT pipeline. As in the actual HERMES processor, an instruction is restarted if a mismatch is detected between the A and B pipelines. The mismatching data is not allowed to commit to the RF, as that would corrupt the processor architectural state. The RF is dual redundant, which allows fast parity protection. The A and B storage is interleaved with approximately 50 µm between redundant nodes. An error is uncorrectable only if the same register and parity group is upset in both copies. The SEU cross-section is diminished over the static tests, presumably due to the higher scrub rate that is achieved in these tests. However, there is a significant window for an SEU to occur after an error is detected by the BIST engine. This leads to detection of events that would not occur in real flux environments observed by spacecraft.

Many of the observed errors are due to SETs in the ALU. These can corrupt all of the data; for instance, a hit on the carry lookahead logic can propagate to all of the adder downstream bits. Erroneous pipeline results can be caught at the write-back stage. These are all corrected before they can corrupt the running processor.

The cache is write-through, and thus there is a copy of this architectural state in the main memory; the primary issue with hardening is to avoid the cache state being corrupted by a soft-error, which would then propagate to the rest of the machine state. The write-through cache architectural state will then naturally be reloaded by the ensuing microprocessor operations. Single cycle invalidation is supported. Particular attention is paid to avoiding SETs, since they could cause upsets such as inadvertent writes or reads of the wrong locations, which in a conventional memory may not be detected by parity or ECC [4,5]. The cache is write-through, and relies on error checking to allow cache invalidation when SEUs or potential SETs are detected. Analysis of the cache proton test results is in progress.

2.   Floating Point Implementation Studies

Any MIPS architecture is free to not include specific instructions, which are then mapped to software. For instance, the Linux kernel has a full FPU emulation inside it. Use of an FP instruction in a machine that does not include an FPU causes an exception. The software then executes the instruction. For this period of the DRDF, we studied simple FPU options for HERMES, focusing on using software instruction emulation. The relative cost of software floating point support was determined by running emulated floating point on the Simplescalar architectural simulator and on the MIPS simulator that we are using for RTL validation; e.g., random and targeted instruction testing. We have looked at putting the Soft-float FPU emulation code into a ROM on the HERMES processor. This approach would allow the code to execute directly from the ROM, with no cache misses and without purging needed instructions from the cache, thus increasing performance about 10× over unassisted software floating point, which would have to fetch FP emulation instructions from main memory and execute from cache. This will still be nearly an order of magnitude slower than with a full FPU, but at very little area and power cost. The code requires 37 kB as compiled with the MIPS gnu compiler. Implemented as a dual redundant memory, which is architecturally simple and compatible with the HERMES

microarchitecture, the required ROM would be less than 0.6 mm^2 in size, fully TID hardened in the target 90 nm IBM fabrication process. This is approximately 3.6× the size of the HERMES RF (less than 2× the size of the FP RF required for a full FPU implementation).

Using the soft-float compiler option and running floating point adds and multiplies through the VMIPS processor simulator, the latter takes 48 and the former takes 86 integer instructions for the emulation. The latter result is surprising, as we expected adds to be less difficult than multiplies. With runs through Simplescalar, FP divides and multiplies use 17× and 68× more clock cycles, respectively. Divide is iterative in any machine, but multiply is pipelined so this is expected. The code compiles differently on Simplescalar and the standard 32 bit MIPS, which accounts for the inconsistency of the FP multiplies. Based on these results, we are now focusing our efforts on a full dual redundant FPU implementation. This will use hardware for all but transcendental functions, which are always implemented in software in the MIPS architecture. The FP RF will be based on the current RF, but upsized to 80 bits (64 plus nibble level parity protection).

3.   HERMES Cryogenic Testing

This aspect of the project will test a HERMES test chip at temperatures from 4 K to 250 K to establish performance of the circuit design techniques in harsh temperature environments. Since the circuits are all-digital, the cold temperature testing should expose any latent race conditions; e.g., hold time issues. Should these appear, greater timing margins will be added to the full HERMES processor circuit design at both synthesis and design timing closure. In this reporting period, effort on this aspect of the DRDF has been on construction of the dipping probe to hold the device under test (DUT) in the dewar containing liquid helium. The probe design is shown both schematically and as fabricated in Figure 3. Figure 4 shows the wide-mouth dewar and the DUT socket. We chose not to use a ZIF socket to avoid moving parts and more mechanical device failure modes at cold temperature. A thermocouple will be directly over the die inside the ceramic package to measure the temperature with less than a 0.5 K error. Wiring the socket to the FPGA board, which is the same as used in TID, ion, and proton beam testing, is in progress. The cryogenic testing will be performed in December 2010.

4.   HERMES Board Trade Study

JPL is designing a printed wiring board with memory (FLASH and DRAM), Ethernet, PCI and serial bus interfaces, as in Figure 5. The board is designed so that it can be used in radiation beam testing, with the plan that flight part replacements can be made with minimal effort for a flight version of the board. Thus, this board will provide a basis for both verification and flight implementation. Using the structure of this board, it will be possible to establish performance per unit power and examine tradeoffs of using different types of memories, EDAC schemes, and bus and system architectures.

In the beam test configuration, the board will have 64 MByte of FLASH, 256 MByte of DDR2 DRAM, Ethernet, PCI and serial interfaces, along with a watchdog timer. All interface and timing functions will be provided by a Xilinx FPGA for the beam board. There will be current test measurement points on the board to measure processor power consumption during

normal operation along with allowing for monitoring of the processor power in support of beam testing. The flight board will change the size of memories being used and will replace the Xilinx FPGA with an RHBD ASIC designed under a future effort.

Currently we are investigating existing flight DRAM options to verify that power savings can be achieved by choosing smaller density DDR2 DRAMs that are acceptable for flight. Error-detection schemes (SEC-DED versus Reed-Soloman) are also being investigated for impacts on reliability and power savings. We are also examining board-level power consumption with an eye towards performance per unit power.

C.   SIGNIFICANCE OF RESULTS

The proton testing showed that the dual redundant checking at WB stage logic is effective at catching upsets to processor architectural state before they are committed to the RF. The RF cross-sections are below the expected values for this technology node (10^{-13} cm^2), but not so low as to put the results in question. Only one TC19 cache test was completed, due to timing problems with the ribbon cable connections to the FPGA controller board. However, that test showed results in line with previous heavy-ion tests of the cache. The TC20 test results are anomalous, with a large number of false detections by the checking circuits. This data is still under analysis, but is inconsistent with previous heavy ion tests.

The lightweight FPU emulation using an onboard ROM does not appear to provide sufficient speedup to provide a compelling FPU approach, when simulated on architectural simulators to determine actual code clock/instruction counts. Consequently, this work is now focused on the microarchitecture of a full dual redundant FPU.

The mechanical design and fabrication for the dipping probe for HERMES cryogenic testing has been completed. Wiring is underway and tests will occur in December 2010.

HERMES will provide very high performance per watt, and with an appropriate instrument or flight computer board it makes possible some types of missions that were not possible in the past. The board trade study is still underway, but preliminary findings indicate that a HERMES board with total power consumption near 1 W is possible. Such a board would help meet the requirements of several low-power mission types, including landers, rovers, net-landers, and aerobots. The board trade study has also identified a useful reinforcing feedback mechanism: as the power for the board decreases, the power required to keep memory contents refreshed decreases. This result and similar ones from the board trade study indicate applicability to some potential follow-on proposals for technology development, including a system designed to enable a very low power "sleep"-like mode for a flight computer.

D.   FINANCIAL STATUS

The total funding for this task was $200,000, of which $166,582 has been expended.

E.   ACKNOWLEDGEMENTS

F.   PUBLICATIONS

None.

G.   REFERENCES

[1]   N. Pettit, N. Hindman, D. Patterson, K. Nielsen, K. Holbert, and L. Clark, "High Speed Redundant Self Correcting Circuits for Radiation Hardened By Design Logic," *Proceedings of RADECS*, 2009: 8 pages.

[2]   X. Yao, D. Patterson, K. Holbert, and L. Clark, "RHBD First Level Embedded Microprocessor Cache Memory," *IEEE Transactions on Nuclear Science* **57** (August 2010): pp. 2089–2097.

[3]   L. Clark, K. Nielsen, and K. Holbert, "Radiation Hardened By Design Digital I/O for High SEE and TID Immunity," *IEEE Transactions on Nuclear Science* **56** (December 2009): pp. 3408–3414.

[4]   P. McDonald, W. Stapor, A. Campbell, and L. Massengill, "Non-Random Single Event Upset Trends," *Transactions on Nuclear Science* **36** (December 1989), pp. 2324–2329.

[5]   L Jacunski, et al., "SEU Immunity: The Effects of Scaling on the Peripheral Circuits of SRAMs," *Transactions on Nuclear Science* **41** (December 1989): pp. 2324–2329.

H.   FIGURES
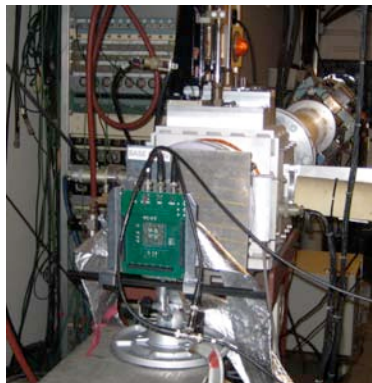
Figure 1. Test apparatus in the LBL beam line. The DUT is in the board in front of the beam as shown.
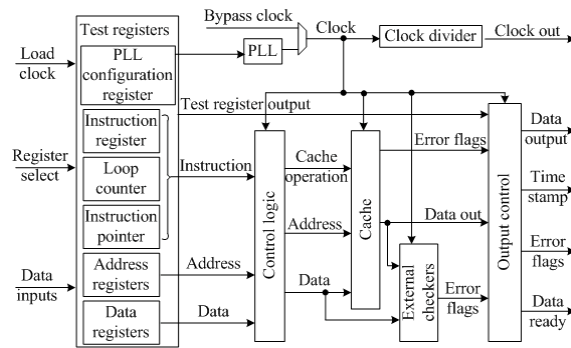
Figure 2. TC19 test chip block diagram showing programmable BIST and cache.
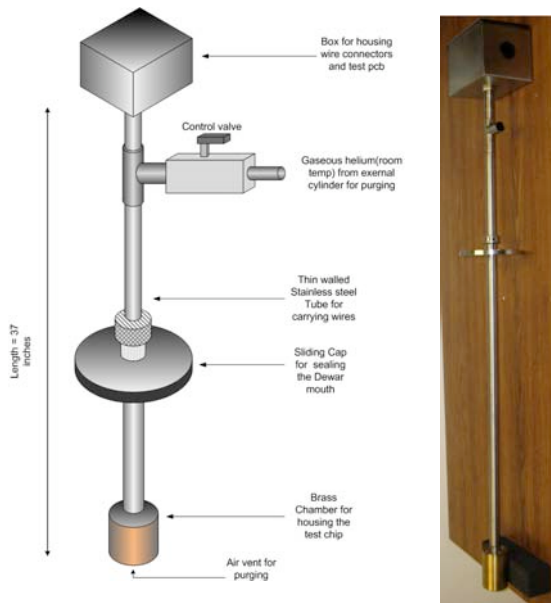
Figure 3. Cryogenic dipping probe schematic view (left) and as fabricated (right).
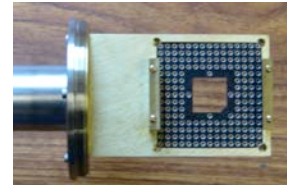
Figure 4. Liquid helium dewar and DUT holder (inset). The DUT holder is threaded to accept the brass chamber to cover the IC during testing.
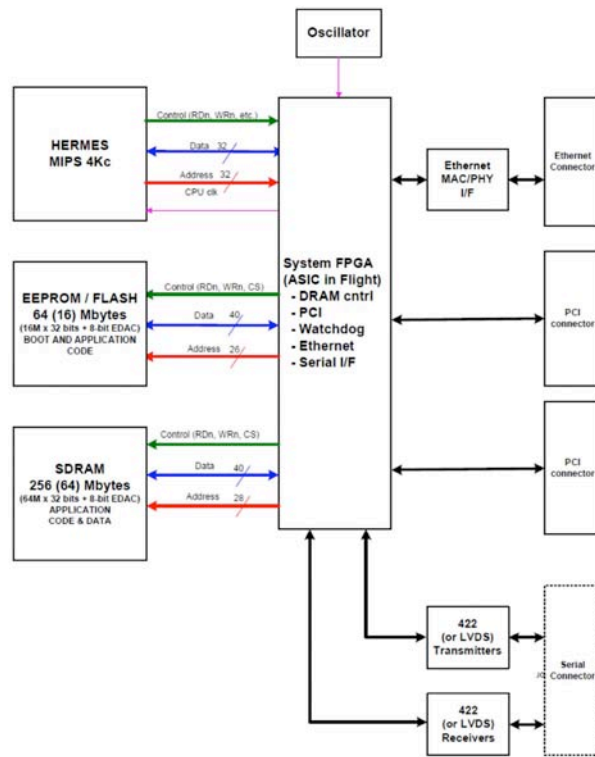
Figure 5. Block diagram for a proposed computer based on HERMES, supporting testing and flight applications.

INFORMATION SYSTEMS

BEYOND MONTE CARLO — STATISTICAL VERIFICATION AND VALIDATION OF SPACE SYSTEMS

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1328

J. Balaram (PI), Mobility and Robotic Systems Section (347)
Raktim Bhattacharya (Co-PI), Aerospace Engineering Department, Texas A&M University
Abhishek Halder, Aerospace Engineering Department, Texas A&M University

A.   OBJECTIVES

Our objective was to develop computationally efficient techniques that go beyond the traditional Monte Carlo (MC) methods of uncertainty analysis in dynamical systems. A considerable body of existing work [1] is directed to using MC methods to obtain the expected value of some function. However, in order to understand the totality of system behavior under non-linear dynamics and non-Gaussian initial probability distribution functions (PDFs), it is useful to evolve the entire PDF of the state and not just try to estimate an expected value and variance. Several statistical analyses can be achieved using moments (mean, variance). For Gaussian distributions, these two quantities are enough to determine the PDF. But the PDF evolution under uncertainty and nonlinear dynamics may not remain Gaussian and higher moments will be needed to approximate the PDF. Accuracy in computation of a given order moment requires exponential increase in the number of sample points required in MC simulations. Thus, our objective was to directly estimate the PDF in a computationally efficient manner. Specifically, we would develop a transfer operation method [2] to evolve the PDF of a dynamical system. Achieving this objective would provide an alternative to Monte Carlo methods to verify and validate future space systems. Our focus was on a demonstration within the entry, descent, and landing (EDL) domain, as that is the one of the more critical phases of a mission, and also represents a computationally intensive use of MC simulation methods.

B.   APPROACH AND RESULTS

We illustrate PDF evolution for a particularly simple dynamical system in Figure 1. A stable one-dimensional linear system: $dy/dt = -a\,y$, has initial states 'y' that are expected to go to the origin with a speed proportional to the magnitude of the parameter 'a'. Our simple example has the initial state 'y' Gaussian distributed with mean of 3 and standard deviation of 1, and the parameter 'a' uniformly distributed in the interval [0.4, 1]. Figure 1(a) shows the PDF at t=0 and Figure 1(b) shows the distribution at t=2. In Figure 1(b), the states for those systems with higher values of 'a', and consequently faster dynamics, get to the origin quicker than those with lower values of 'a'. The skewed shape of the distribution in Figure 1(b) reflects this difference in behavior over the range of 'a'. We now discuss various approaches to determining this PDF evolution for a complex dynamical system. The limitations of MC-based dispersion analysis are well known; i.e., poor computational scalability and lack of a methodology to quantify the

evolution of uncertainty in a statistically consistent manner. Consequently, researchers have pursued different methods for uncertainty propagation in dynamical systems. The methods may be grouped in two broad categories: parametric (where one evolves the statistical moments) and non-parametric (where one evolves the full PDF).

There have been three major directions in parametric propagation of uncertainty:

- The simplest method in this category assumes a linear state space description of the form $dx/dt = A\, x$. With the initial condition uncertainty assumed to be Gaussian, one then propagates the mean and covariance matrix by the well-known state and Riccatti equations. The major drawback is that a typical EDL system with large number of states involving highly nonlinear dynamics and non-Gaussian initial joint PDF does not fit in this framework.

- In polynomial chaos (PC) method, one derives a set of deterministic ordinary differential equations (ODEs) using either Galerkin projection or stochastic collocation and then solves the resulting set of ODEs. Although this method can handle nonlinear dynamics with non-Gaussian uncertainties, one ends up solving a higher-dimensional state-space problem, which becomes intractable for a realistic EDL simulation. Further, the method is difficult to apply for large nonlinearities and the computational performance degrades (due to finite-dimensional approximation of the probability space) if long-term statistics are desired. A recent attempt to apply this method in EDL domain can be found in [3].

- Another method in this category is called the direct quadrature method of moments, where the PDF is approximated as a sum of Dirac delta functions with evolving parameters. This method has a problem of meeting various closure conditions on the resulting moments.

Non-parametric propagation of uncertainties can be performed using either an approximate method or a direct method.

- In the approximate method, one tries to estimate (in non-parametric sense) the underlying PDF. The method aims to approximate the solution of the PDF transport equation. This method is widely exercised in statistics community [4] under the name of kernel density estimation, although most applications there a concerned with static data. In a dynamical system, one must do a dynamic optimization to determine the optimal values of the parameters at every time instant. The method can suffer from high computational cost arising due to the explicit enforcement of normality constraint and moment closure constraint at each step of dynamic optimization procedure. Moreover, for high-dimensional state spaces like planetary EDL, recursively performing constrained optimization becomes challenging.

- In the direct method, one works to directly solve the PDF transport equation and as opposed to approximating its solution. If the dynamics are assumed to be deterministic, this transport equation reduces to the stochastic Liouville equation (SLE) [5], which is a quasi-linear partial differential equation (PDE), first order in both space and time. The SLE is an example of the Perron-Frobenius (PF) operator [2] method of propagating the PDF

evolution. The evolution of the PDF is given by the action of the PF operator P_t on the density $f(y)$; i.e., $u(t,y) = P_t f(y)$, which in turn satisfies the partial differential equation $\partial u/\partial t + \sum \partial(uF_i)/\partial y_i = 0$, where F represents the dynamics of the system and the summation is over all the state components. The SLE can be easily solved in a direct way using the method of characteristics (MOC). Since all the statistics can be derived from the PDF, from an information point of view, it is definitely superior to the parametric propagation methods. In the SLE method, instead of individual realizations (initial conditions and/or parameters), one propagates the ensemble of realizations. This, in essence, means that the number density of trajectories meets the continuum hypothesis [5]. Since the SLE is solved directly, the solutions will satisfy the criteria to be a PDF, and hence the problems like moment closure or normality constraints are not required to be explicitly enforced.

We had originally considered also analyzing solution methods for dynamical systems under conditions of stochastic forcing; i.e., solving transport equations using the Fokker-Plank-Kolmogorov (FPK) formulation [6]. However, we have since chosen to focus on solving the PF operator-derived SLE equations. This allowed us to give a more in-depth treatment to the SLE techniques for PF solutions and defer the FP formulation to a later effort.

Milestone. PF Solver Implementation — We developed PF solvers based upon solving the SLE equation in two different implementations utilizing Matlab and C++ codes for systems specified by analytical equations for the system dynamics. In this case, the gradient terms needed for the SLE solution can be obtained analytically. This was applied to the EDL problem to solve the so-called Vinh's equations [7]. We used the evolved PDFs to compute the marginal distributions that correspond to those typically used in EDL analysis. The results of using this solver and the corresponding MC results are shown in Figure 2 for the case of the longitudinal Vinh's equations with three states, height (h), velocity (V) and flight-path angle (FPA) solved for 5000 initial conditions. On the left, the red (solid) uni-variate curve corresponds to PF and the blue (dash-dot) lines to MC. The MC and PF uni-variate densities are in good match. To the right are the bi-variate marginal distributions, with the top row in the bi-variate plots corresponding to MC and the bottom one to PF solutions, respectively. The bi-variate marginals show the general trend that PF-derived marginals capture the concentration of probability mass well (by virtue of the probability weights obtained by solving SLE), while the MC results tend to smear it out (because of bin-counting). This can be seen, for example, in FPA vs. V bi-variate plots.

Milestones. Scaling Software for handling large number of states; Code module testing at JPL for example EDL application — For systems with a large number of states (e.g., systems with attitude dynamics or multiple bodies), we added a PF solver using the MOC to the DSENDS simulation system at JPL. DSENDS uses a computationally efficient multibody dynamics engine with dynamics computation scaling linearly with the degree of freedom. Our initial implementation uses a numerical differencing method to compute the gradient terms needed for the SLE solver. We then applied this newly extended DSENDS system to a realistic example involving the Mars Pathfinder spacecraft together with high-fidelity atmospheric and nonlinear aerodynamics models. We show some of the numerical results from these simulations. The graph on the left in Figure 3 shows the evolution of the probability weights for a system with both translational and attitude dynamics within a single solution (a 12-state system) obtained by the MOC. It can be seen that, in the initial portion of the trajectory, the weight peaks as the

system slows down and the trajectories crowd into a smaller region of the velocity space. However, towards the latter part of the trajectory, the weight starts to decrease to reflect the dispersion produced by the attitude instability in that portion of the flight regime. The plot on the right in Figure 3 shows the evolution of solutions obtained by the MOC for an ensemble of trajectories. The multidimensional state has been projected into relative-flight-path (rfpa) and atmosphere-relative velocity magnitude (rmag) coordinates. The individual solutions are color-coded with their end-of-trajectory probability weights, with a rainbow color coding in which violet-blue reflect lower values and orange-red reflect larger values.

Milestone: Benchmarking performance and verification — A specific quantitative objective was to determine the PDF of the system state for systems with dozens of states with parametric uncertainty in 100s of parameters using 2 orders of magnitude (i.e., 100×) less computation than the traditional Monte Carlo methods. We had shown this computational advantage for simple systems with benchmarking results in the initial proposal. However, during the performance of the DRDF task, we came to the realization that this benchmarking was more difficult than we had anticipated for the more complex systems. The PDF solution using the MOC gives the exact probability weight for the evolved state PDF for a particular point in the space with far less computation than an MC histogram method of estimating the PDF around that point. However, in most EDL analysis, the MC method is not used to directly obtain PDF values but is used instead to obtain estimates of various conditional distributions of the state — e.g., distribution of the landing footprint; i.e., P(x,y|z=0). Obtaining these conditional distributions from a PDF requires evaluating the associated marginal distributions for the evolved PDF, which in turn are obtained as integrals over the PDF. Note that when computing integrals (i.e., expectations), whether one uses the evolved state with its distribution (obtained by the solving SLE using the MOC) or the initial distribution, in the limit where a large number of points are sampled, the results are mathematically equivalent. One therefore has to compare the computational cost of obtaining the marginal distribution from the evolved PDF using a limited set of sample points to those obtained by MC estimates obtained by histograms generated from the initial PDF distributions. An additional factor that has to be resolved in the benchmarking process is to hold the desired accuracy level for each method to be the same. We are currently performing computational experiments using the low-dimensional forms of the analytical Vinh's equations to address these questions.

Specifically, we are:

• Using the Kullback Liebler (KL) measure to determine how close the numerically obtained distributions are to the truth distribution (where computable analytically).

• Investigating integration methods in the high-dimensional spaces to obtain the "area" under the evolved PDFs, and techniques such as adaptive refinement to ensure that the evolved PDF values are obtained over the region of interest.

• Determining cases where it may be advantageous to obtain conditional distributions using the evolved PDF instead of traditional MC estimation of the integrals.

C. SIGNIFICANCE OF RESULTS

The techniques we have demonstrated to date give us an exact method to obtain the PDF as evolved through a dynamical system. Using the PF method, if one knows the functional form of the initial PDF, one could generate any desired number of samples of that PDF using Markov Chain Monte Carlo (MCMC) and then propagate the PDF for these "high-initial-probability" samples. The samples would be assigned weights (value of the joint PDF) at later instants of time by solving the Liouville equation using the method of characteristics. Since the exact probability weights are calculated in PF method, one no longer needs huge number of samples to approximate the PDF. On the contrary, in the MC method, approximate joint PDF construction at a given time instant requires binning and counting samples in a large number of histogram bins.

The temporal evolution of the PDF gives the EDL analyst additional insight into the behavior of the system under uncertainty. Especially in the case of low-dimensional EDL systems, the numerical results appear to be more precise when compared to some of the "smeared out" MC results. An initial implementation of MOC solution methodology is now available in the DSENDS simulator system and allows for high-fidelity numerical experiments.

For statistical analyses, which are based on expectation and variance, the PF approach may not provide a significant advantage over MC in terms of accuracy and computational time. As discussed earlier, for higher-dimensional systems where conditional distributions (using expectations) needs to be determined, the direct advantage of the PF methodology has not yet been established. Nevertheless, we are currently evaluating if the PF solutions can be used in conjunction with other MC methods such as importance sampling or MCMC, where various sampler and proposal densities need to be determined and could be obtained from the PF solutions. However, for analyses where the PDF is required, PF will outperform MC. Examples of such analyses include state and parameter estimation, robust and optimal control with probabilistic uncertainty, and system reliability analysis where system-level uncertainty is determined based on component-level uncertainty.

D. NEW TECHNOLOGY

The emphasis here has been on algorithmic exploration of uncertainty propagation in dynamical systems. The implementation of the MOC method for evolving the PDF represents a new software technology now available as an initial implementation in the DSENDS systems. After further validation and refinement, this will be reported as a new software technology.

E. FINANCIAL STATUS

The total funding for this task was $200,000, all of which has been expended.

F. ACKNOWLEDGEMENTS

The PI would like to thank Aron Wolf, who helped motivate the EDL problems and was on the original proposal as a Co-I. The PI would also like to thank David Bayard and Abhi Jain for many insightful discussions on MC methods and dynamical systems. The support of the Mars

G. PUBLICATIONS AND PRESENTATIONS

[A] Abhishek Halder and Raktim Bhattacharya, "Beyond Monte Carlo: A Computational Framework for Uncertainty Propagation in Planetary Entry, Descent and Landing," submitted to *AIAA GNC Conference,* Toronto, Canada, 2010.

H. REFERENCES

[1] Byron J. T. Morgan, *Elements of Simulation*, CRC Press, 1984.

[2] A. Lasota and M.C. Mackey, "Chaos, Fractals and Noise," *Stochastic Aspects of Dynamics*, Springer, 1994.

[3] A. Prabhakar, J. Fisher, and R. Bhattacharya, "Polynomial Chaos Based Analysis of Probabilistic Uncertainty in Hypersonic Flight Dynamics," *Journal of Guidance, Control and Dynamics* **33** (2010): pp. 222–234.

[4] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*, Monographs on Statistics and Applied Probability, 26, Chapman & Hall, 1994.

[5] M. Ehrendorfer, "The Liouville Equation and its Potential Usefulness for the Prediction of Forecast Skill. Part I: Theory," *Monthly Weather Review* **122** (1994): pp.703–713.

[6] M. Kumar, S. Chakravorty, and J. L. Junkins, "A Homotopic Approach to Domain Determination and Solution Refinement for the Fokker-Planck Equation," *Probabilistic Engineering Mechanics*, in press.

[7] N. X. Vinh, A. Busemann, and R. D. Culp, *Hypersonic and Planetary Entry Flight Mechanics*, University of Michigan Press, Ann Arbor, 1980.

I. FIGURES
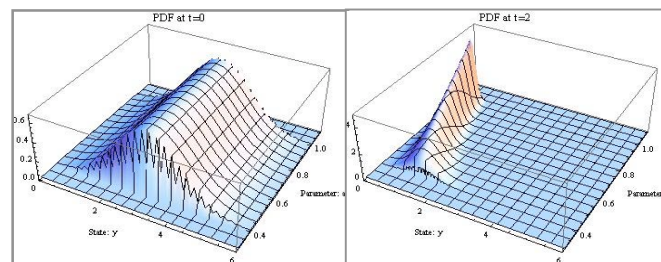
Figure 1. Example of PDF evolution for a simple linear system.
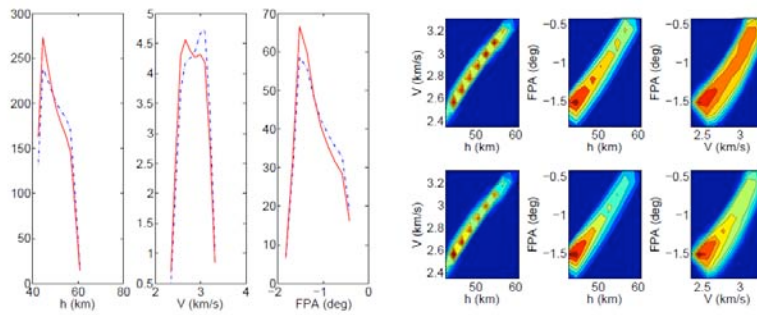
Figure 2. Solutions obtained by both MC and PF methods.
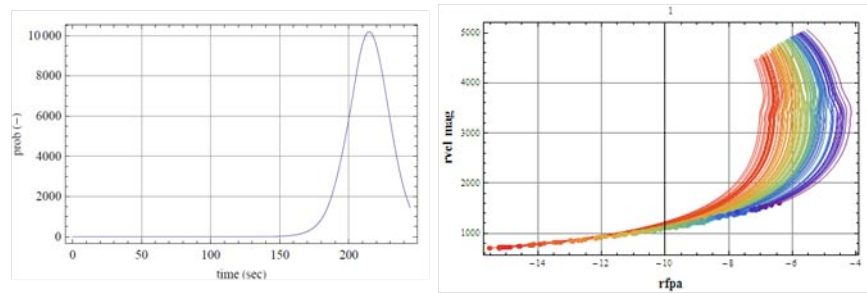
Figure 3. Evolution of PDF weights and ensemble of solutions.

MATERIALS

ACOUSTIC BANDGAP-BASED SHOCK/VIBRATION PROTECTION SYSTEM

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1321

R. Peter Dillon (PI), Electronic Packaging and Fabrication Engineering Section (374)
Gregory L. Davis (Co-PI), Mechanical Systems Engineering (350)
Chiara Daraio (Co-PI), Aeronautics and Applied Physics, California Institute of Technology
Nicholas Boechler, Aeronautics, California Institute of Technology

A.   OBJECTIVES

The objective of the proposed research was to develop and build a tunable acoustic bandgap material for use in a shock/vibration protection system, demonstrating TRL 4–5 using well-established theoretical, numerical, and experimental techniques at Caltech and at JPL's extensive environmental test facility. The deliverables/products of the work are a report on the design, manufacture, and test of a multiparticle periodic array, and an assembled and tested breadboard array [1].

B.   APPROACH AND RESULTS

The research undertaken and accomplished is outlined by the following major milestones: define shock and vibration requirements; design a multiperiodic array with a bandgap tunable to the required spectrum by precompression; build and package an array; test the array with a model assembly for performance and efficiency; and analyze the test data and report results.

The following requirements were used to frame the design of an acoustic bandgap-based shock/vibration protection system for a model spacecraft assembly. The system will show an order of magnitude reduction in the amplitude of the shock response spectrum (SRS), Figure 1, for frequencies above 1 kHz. The preloaded, multiparticle, periodic array being designed for shock/vibration filtering will be suitable for system and payload scale assemblies [1].

Given these constraints, a multiparticle periodic array with a bandgap tunable to a cutoff frequency of 1 kHz was designed at Caltech utilizing experimental, theoretical, and computational methods. The experimental setup, Figure 2, consisted of a stack of spheres held together by polycarbonate framing. Dynamic perturbations were introduced by either a piezo stack actuator or an electrodynamic shaker. Piezo sensors were used to measure wave propagation at various positions in the stack. An example of force-time measurements (and the vibration protection provided by a dispersive filter) from this setup is shown in Figure 3. While a 500-Hz wave propagates through the chain, the force output at 1169 Hz, a frequency in the band gap, is largely filtered. There is also significant agreement between the experimental and computational results, as shown in Figure 4. Scalability of the bandgap based on materials selection and particle size are shown experimentally and computationally in Figure 5. The materials and particles' sizes investigated are listed in Table 1. Their performance has been

experimentally measured and the results for 1.5" steel and polyurethane spheres are shown in Figure 6 [1]. Polyurethane exhibited a cutoff frequency of 1 kHz, which coincided with the established requirements and consequently was chosen as the multiparticle array material.

For functional deployment of a multiparticle periodic array of spheres for shock/vibration protection, an in situ method for precompressing the particle stack was required. Additionally, the input and output ends of the array needed to couple with the excitation source and the sensitive payload, respectively. The array also needed to be free to move within its container and the container needed to be decoupled from the excitation source. To achieve this, the container was machined from a metal cylinder to provide clearance for a stack of polyurethane spheres. Enough clearance was provided to accommodate the spheres' dilated diameter under precompression. Pistons at each end of the cylinder make contact with the linear array of spheres. Precompression of the array is achieved through the use of springs located between the pistons and end caps, which screw onto the container. Compression of the springs by tightening the end caps allow the amount of precompression to be controlled. The threaded ends of the piston rods extend through Teflon linear bearings which are press-fit into the end caps. A cartoon of this concept is illustrated in Figure 7. A picture of the fabricated and assembled prototype is shown in Figure 8a.

Pyroshock testing of the prototype was performed on the Environmental Test Lab's tunable beam, Figure 8b. The shock environment was monitored in terms of a shock response spectrum (SRS), Figure 9a, by accelerometers mounted on an interface plate at the input side of the prototype and on 1.67- and 10-kg mass models at the output, "protection" side of the prototype. The applied shocks were progressively increased by increasing the impactor launch pressure from 50 psi to 200 psi. Above the cutoff frequency, shock attenuation was at least 99%, a reduction of 2 orders of magnitude, for all impactor launch pressures, with the higher pressures appearing to exhibit more idealized band gap behavior. At 200 psi, the cutoff frequency appears to occur at ~800 Hz, Figure 9b. Shock attenuation by 1 order of magnitude was expected as a result of the dispersive effects of the linear array. Additional damping is contributed synergistically by the structural mass of the prototype and dissipative effects as evidenced by the ~90%, single-order-of-magnitude attenuation at frequencies in the passband. The material damping contribution is confirmed by replacing the polyurethane spheres with a solid polyurethane rod in a shaker test, Figure 10. Attenuation is ~2 orders of magnitude in power spectral density (PSD), which equates to ~1 order of magnitude in the shock response spectrum.

C.   SIGNIFICANCE OF RESULTS

A prototype shock/vibe protection system that leverages the potential of phononic acoustic band gaps was designed, built, and tested. A cutoff frequency of ≤1 kHz was demonstrated with significantly more than 90% attenuation. The prototype reveals a synergistic bandgap/damping effect when packaging linear arrays for shock/vibe protection in the method demonstrated. The prototype is configured as to be deployable in a hexapod configuration for system or payload protection.

While the preliminary results cited in our proposal described a dimer system with a bandgap between 7 and 14 kHz [2], our results have shown a variety of acoustic bandgaps,

ranging in width and frequency, and cutoff frequencies are attainable through the proper selection of materials, size, and stacking. Attenuation in the band gap or above the cutoff frequency shows very little dependence on the magnitude of the input excitation. This technology has significant potential as mission enabling.

D.   NEW TECHNOLOGY

A New Technology Report and Patent Application are in preparation.

E.   FINANCIAL STATUS

F.   ACKNOWLEDGEMENTS

G.   PUBLICATIONS AND PRESENTATIONS

[A]   N. Boechler, R. P. Dillon, J. P. Borgonia, D. L. Kahn, A. A. Shapiro, G. L. Davis, and C. Daraio, "A Granular Media-Based Tunable Passive Vibration Suppressor with Self-Contained Mechanism for the Application of Static Compression," *Journal of Acoustical Society of America* (in preparation).

[B]   N. Boechler, R. P. Dillon, J. P. Borgonia, D. L. Kahn, A. A. Shapiro, G. L. Davis, and C. Daraio, presentation in preparation for Engineering Mechanics Institute, Los Angeles, California, August 2010.

H.   REFERENCES

[1]   R. P. Dillon, G. L. Davis, C. Daraio, and N. Boechler, "Acoustic Bandgap-Based Shock/Vibration Protection System," DRDF Interim Annual Report, 2009.

[2]   E. B. Herbold, J. Kim, V. F. Nesterenko, S. Wang, and C. Daraio, "Tunable Frequency Bandgap and Pulse Propagation in a Strongly Nonlinear Diatomic Chain," *Physical Review E* (in review), 2008.

I. FIGURES
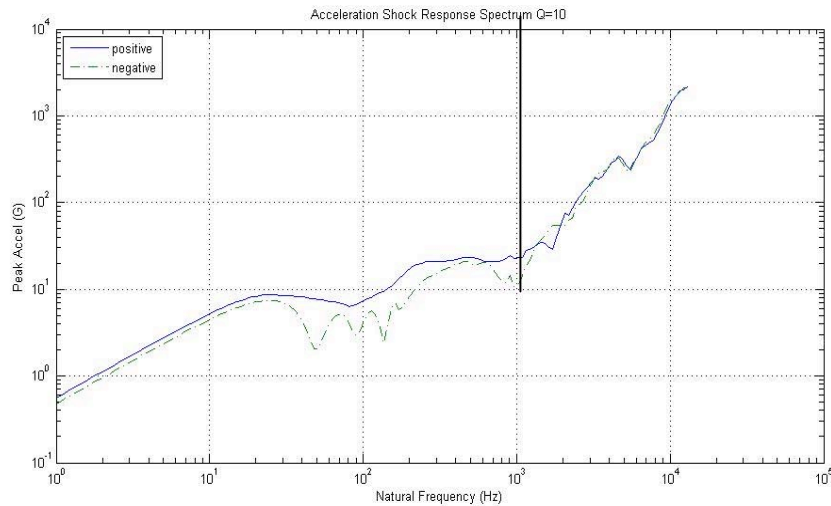
Figure 1. Example shock response spectrum (SRS). The shock vibration protection system designed reduces peak accelerations above 1 kHz by more than an order of magnitude.

(a)   (b)
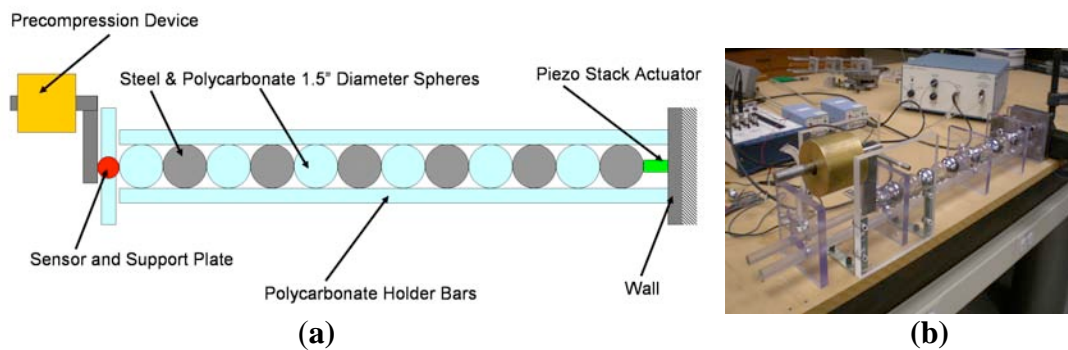
Figure 2. A cartoon (a) and photo (b) of the experimental setup. Not shown are the piezo sensors used to measure wave propagation in the stack. They may be substituted for any sphere in the stack to determine the wave characteristics as a function of stack position.
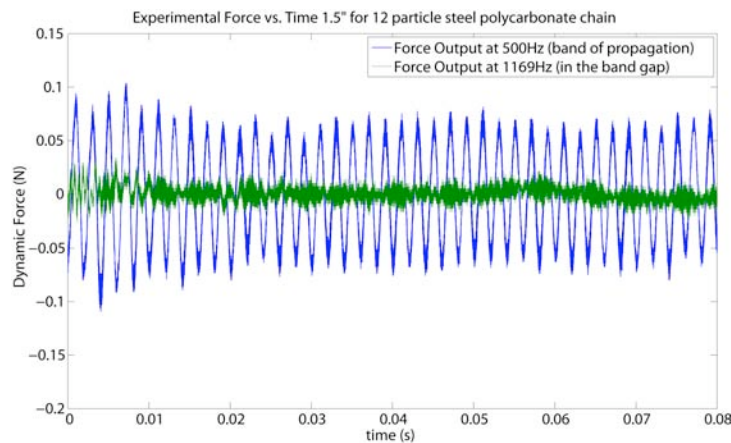
Figure 3. Force-time measurements for alternating steel and polycarbonate beads.
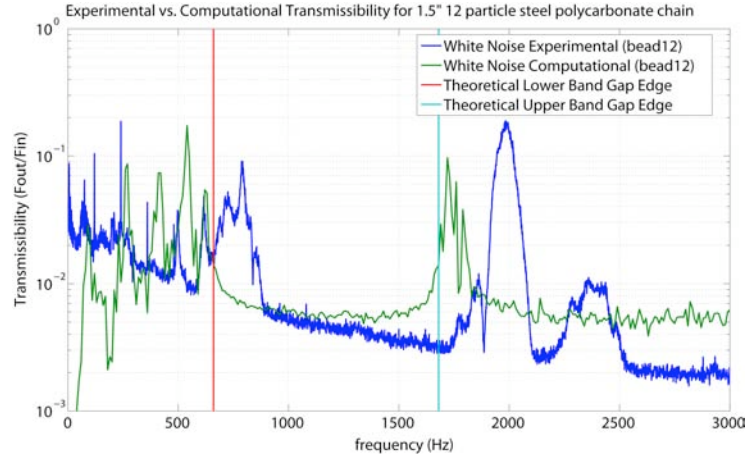
Figure 4. Experimental vs. computational bandgap for alternating steel and polycarbonate beads.

(a)   **(b)**
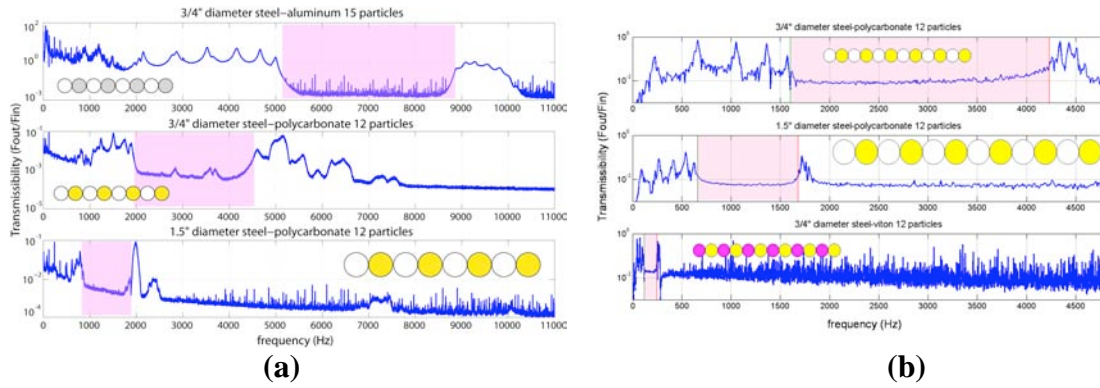
Figure 5. Experimental (a) and computational (b) scalability results for varying materials and sizes.
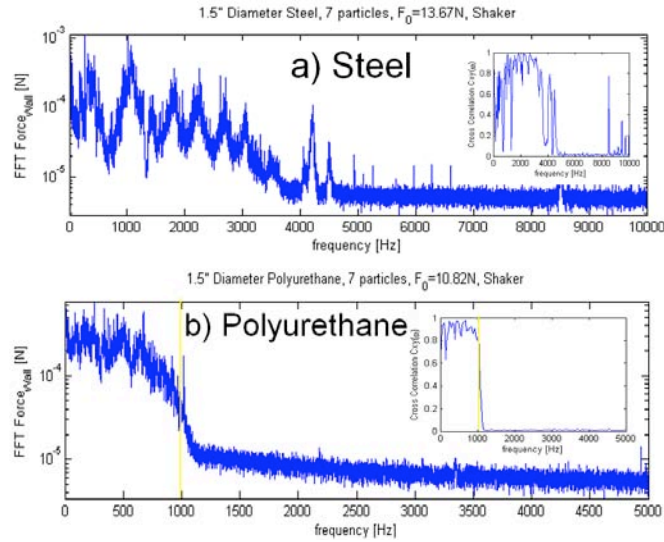
Figure 6. Experimental bandgap results for varying 1.5" steel (a) and polyurethane (b) spheres.
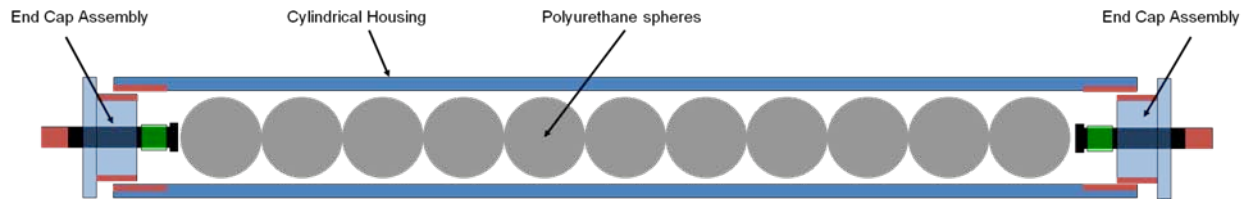
Figure 7. Phonon bandgap-based shock/vibration protection system concept.
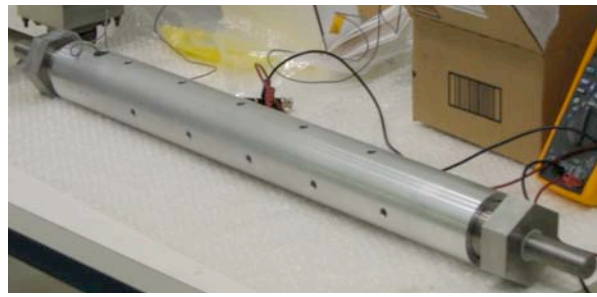
(a)   (b)

Figure 8. Phonon bandgap-based shock/vibration protection system prototype assembled (a) and on the Environmental Test Lab's tunable beam with mass model (b).
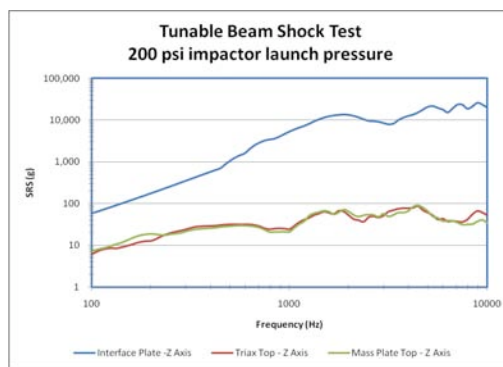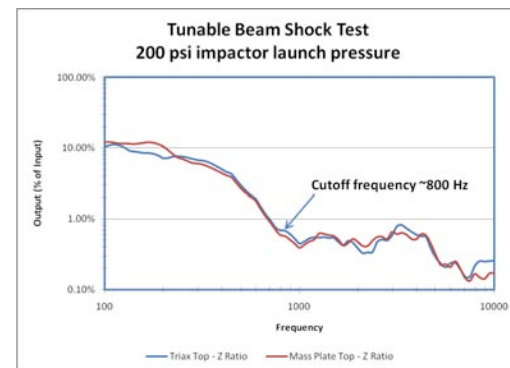
(a)   (b)

Figure 9. (a) Shock response spectra (SRS) from accelerometer data at the interface plate and the mass model of a 200 psi shock test; (b) The ratio of the output SRS data to the input SRS data.
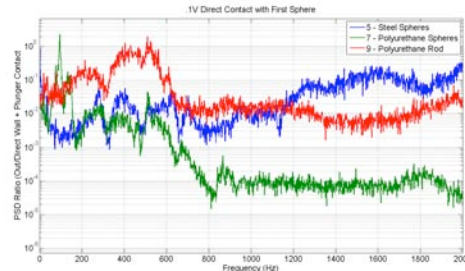
Figure 10. Power spectral density of the prototype with steel spheres (blue), polyurethane spheres (green), and a solid polyurethane rod (red).

Table 1. Acoustic bandgap materials and sizes investigated.

#	Material	Diameter (in.)	Shape
1	316 Stainless Steel	0.75	Sphere
2	6061 Aluminum	0.75	Sphere
3	316 Stainless Steel	0.75	Cylinder (3/4" length)
4	316 Stainless Steel	0.75	Cylinder (3/8" length)
5	316 Stainless Steel	0.75	Cylinder (1/4" length)
6	316 Stainless Steel	1.5	Cylinder (1.5" length)
7	Viton Rubber	0.75	Sphere
8	E52100 Steel	1.5	Sphere
9	Polycarbonate	1.5	Sphere
10	Polycarbonate	0.75	Sphere
11	Nylon	0.75	Sphere
12	Buna-N	0.75	Sphere
13	PTFE	0.75	Sphere
14	Polyurethane (95 Shore A)	0.75	Sphere
15	Polyurethane (95 Shore A)	1.5	Sphere
16	Polyurethane (95 Shore A)	3	Sphere
17	E52100 Steel	3	Sphere
18	Polyurethane (95 Shore A)	1.5	Sphere (w/1.25" steel core)
19	Polyurethane (95 Shore A)	2	Sphere

MICRO/NANO DEVICES

HIGH-PERFORMANCE 45-NANOMETER SILICON-ON-INSULATOR COMPLEMENTARY METAL OXIDE SEMICONDUCTOR TECHNOLOGY FOR ARRAY REMOTE SENSING, RADAR, AND COMMUNICATION SYSTEMS

Director's Research and Development Fund (DRDF)
Interim Report

JPL Task #1352

Andy Fung (PI), Instrument Electronics and Sensors Section (389)
Lorene Samoska, Instrument Electronics and Sensors Section (389)
Erich Schlecht, Instrument Electronics and Sensors Section (389)
Todd Gaier, Instrument Systems Implementation and Concepts Section (382)
Gabriel Rebeiz, Department of Electrical and Computer Engineering,
University of California at San Diego

A.   OBJECTIVES

Our objective is to evaluate the operating frequency of state-of-the-art 45-nm silicon-on-insulator (SOI) complementary metal oxide semiconductor (CMOS) technology. We do this by pushing the high-frequency limit of the technology for circuits targeting operation up to 220 GHz for applications in remote sensing, radar, and communication systems. This study will provide information on how the technology can be utilized for analog electronics in future programs and missions. As an evaluation vehicle, the goals are to design and test various high-frequency circuit components towards developing electronics for an 180-GHz radiometer integrated circuit (IC) element applicable for an array system. First-iteration circuit designs that have been fabricated and tested are an 85 to 105 GHz low-noise amplifier (LNA), 160 to 220 GHz LNA, 160 to 220 GHz single-pole-double-throw (SPDT) switch, 160 to 220 GHz frequency doubler, and 180 GHz detector diode circuit, all in a 45-nm SOI CMOS technology. Some of these circuit types are the first Si-based circuits designed and fabricated targeting operation up to 220 GHz. It should be noted that it is well known that optimized III-V compound semiconductor transistors and cooled superconductor-based devices can demonstrate higher performance capabilities than Si CMOS. The goal of this effort is not to demonstrate necessarily superior SOI CMOS circuit performance but to determine the highest frequency circuit capabilities of the latest 45-nm SOI CMOS technology, and how the technology may be used to implement new high-frequency, highly integrated electronic systems where other technologies cannot, due to their limited integration yield. The strength of CMOS is in its ultra-large-scale integration capability coupled with increasing operating speeds and reduction of power consumption from transistor-size scaling. For least expensive mass production of complex, highly integrated low-power IC systems, Si CMOS is arguably the best technology. SOI CMOS also has benefits of being more immune to radiation effects and not prone to latch-up like bulk Si CMOS technology.

B.   APPROACH AND RESULTS

Tasks in this proposal are achieved through the following methods and efforts.

Circuit design method. 1) The IBM 45-nm SOI process results in a CMOS transistor with an ft/fmax of 385/480 GHz, and is one of the highest available in any technology even when compared to standard 32-nm bulk silicon CMOS. The transistors are characterized by IBM at low frequencies for developing their design kits. In this effort, the low-frequency circuit models are used to design circuits up to 220 GHz. The passive components (L, C, and transmission lines) are all modeled using Sonnet and HFSS. Circuit designs are done using CADENCE and follow the standard design flow of LNAs, SPDT switches, frequency doublers, etc. (impedance matching, nonlinear frequency generation, filtering, and DC response/responsivity).

Circuit layout. The circuits are laid out manually using the IBM 45-nm SOI guidelines, which have stringent metal density rules, and then resimulated using HFSS and Sonnet to take into account the additional metal layers and islands placed next to the circuit required by design rules. All designs have to pass the Design Rule Checker (DRC) and Layout versus Schematic (LVS) checker before being accepted by IBM for fabrication.

Fabrication. The circuits are fabricated in a multiple-project wafer run (MPW) on 8-inch SOI wafers. There are 11 layers of wiring above the SOI transistor, for digital routing, bias supply routing, and RF routing. The wafer run time is approximately 4 months, and is available 2 to 3 times per year. DARPA supports the fabrication costs and grants UCSD generous space on the SOI wafer (30 mm^2).

Testing. For characterizing circuits, S-parameters, noise figure (NF), power-in versus power-out (PiPo), detector noise, and detector responsivity are measured where appropriate. Circuits are first examined for DC and RF functionality through DC bias and S-parameter checks. S-parameters are performed on-wafer in two frequency bands; they are the 90–140 GHz (WR8) band and the 140–220 GHz (WR5) band. NF can be measured from 75–110 GHz (WR10) band with a Quinstar noise source and or through the Y-factor method with 77 K and room temperature load sources at higher frequencies. Input power versus output power (PiPo) response for the frequency doubler can be measured using a 90-GHz, frequency-multiplied, calibrated-power input source, with the frequency-doubled output power measured with a calorimeter. A mixer downconverter with spectrum analyzer can be used to verify the output frequency. Detector circuit RMS noise voltage and responsivity can be measured with a spectrum analyzer at the circuit output, while the circuit is provided with an input matched load and a calibrated input power source, respectively.

Tasks and Milestones.

1)   Simulate and study performance of first iteration circuits (UCSD) [February 2010 to March 2010]

2)   First-iteration chips arrive from IBM foundry (milestone) [March 15, 2010]

3)   Setup and test where appropriate: dc bias, S-parameters, frequency multiplication, noise figure, noise voltage, and responsivity (JPL and UCSD) [March 2010 to July 2010]

4)   Evaluate measurements against simulations and determine model/circuit tweaks required (UCSD) [April 2010 to July 2010]

5)   Design second-iteration circuit improvements and IC chips, and mixers (UCSD) [April 2010 to July 2010]

6)   Second-iteration chips arrive from IBM foundry (milestone) [December 2010]

7)   Selected circuit characterization and analysis (JPL and UCSD) [January 2011 to February 2011/program end]. This will be similar to Tasks 2) through 5) above.

We have performed items 1) through 5) above and second-iteration chips are now in fabrication at IBM, with anticipated delivery in December 2010. First-iteration chips were designed based on the available low-frequency IBM device model and design kit, and electromagnetic simulations of passive components. After fabrication of the first-iteration designs, the circuits were characterized. For amplifier designs (1-stage common source, 3-stage common source, 4-stage common source, 1-stage cascade, 4-stage cascade), S-parameter measurements from 75–110 GHz and 140–220 GHz indicate their frequency responses are tuned low and their gains to be less than that from simulations. Figure 1(a) shows the response of a 3-stage common-source amplifier and Figure 1(b) shows a 1-stage cascade amplifier. For the 3-stage common-source amplifier, the amplifier had sufficient gain performance so that NF data were also taken. For this amplifier, its lowest NF was measured to be 5.6 dB at 85 GHz. The optimum operating current density for the lowest NF is 0.225 mA/μm. Current is normalized to transistor gate width. Figure 2 shows the measured NF data. Single-pole single-throw (SPST) and single-pole double-throw (SPDT) S-parameter measurements show that the insertion loss for the on-state of switches is about 2 to 6 dB greater than simulations, and off-state isolation to be about 10 db less than simulations. On-state return loss (port match) is typically within 5 dB of simulations for the switches. Figure 3 shows S-parameters of the SPST switch. Performance of the frequency doubler was also tuned lower in output return loss (see Figure 4). Its conversion loss at 170 and 180 GHz is around 20 to 25 dB. S-parameter measurements also show the detector diode circuit to be matched lower in frequency and no additional measurements were made on it.

For first-iteration circuits, differences between simulations and measurements are attributed to device model error, unaccounted parasitic effects, and bias line coupling effects that caused oscillations. The foundry device models that we used as a starting point are developed for much lower frequency simulations and are not accurate enough to be directly used for 220-GHz circuits. The device models did not include the routing parasitics between transistor fingers, such as metal to polysilicon overlap capacitance, via inductance and material sheet resistances. Also, measurements of passive test structures in the first iteration indicated that the transmission line wiring used in circuits had characteristics that differed from simulations. Measurements and subsequent analysis also indicated the bias lines and power supply bus can resonate with capacitors. To compensate for these effects in the second-iteration circuits, the device model has been augmented with extracted device routing parasitic. The interconnect stack up has been updated mainly through the wiring's effective dielectric constant. Plus on-chip capacitors and resistors have been added to allow for power supply decoupling from the chip and detuning of simulated oscillations. Figure 5 shows some of the new circuits in the second iteration presently in fabrication.

C.   SIGNIFICANCE OF RESULTS

We have designed and tested some of the first 45-nm SOI CMOS circuits targeting operation up to 220 GHz. Although measured performance differed from simulations, the findings have allowed us to determine initial capabilities of this technology and how the baseline design method applies to higher frequencies. We have identified some causes of the differences between measurements and simulations and applied changes to the design method to compensate. A second set of circuits with the changes is in fabrication presently and will be measured when the circuits are available. In an iterative approach, we are improving our understanding of the capabilities of 45-nm SOI CMOS technology and how to better use it.

The knowledge we are gaining and the modifications we are applying to the baseline design method is advancing the technology and will allow us to better utilize high-yield silicon CMOS traditionally used at lower frequencies for much higher frequency applications in NASA/JPL programs and missions. Uses can include implementation of electronic array systems for remote sensing, radar, and communication transceivers.

D.   FINANCIAL STATUS

The total funding for this task was $180,000, of which $124,182 has been expended.

E.   ACKNOWLEDGEMENTS

University of California San Diego graduate students Woorim Shin, Mehmet Uzunkol, and Ozgur Inac have contributed significantly to this effort.

F.   PUBLICATIONS

None.

G.   REFERENCES

None.
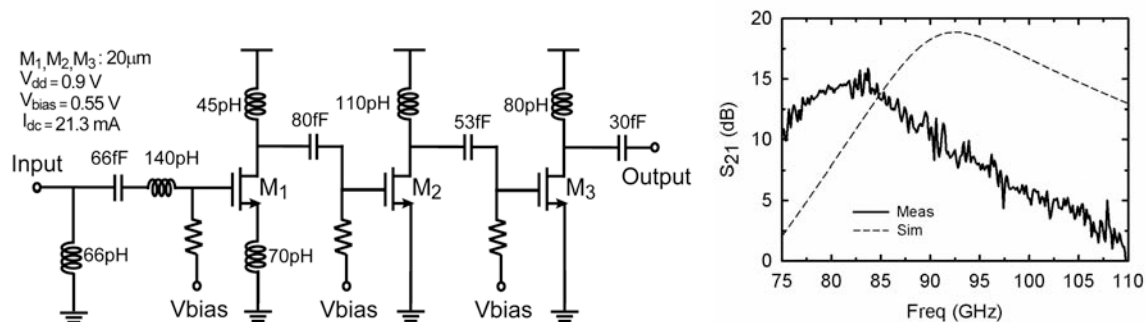
H.   FIGURES

Figure 1 (a). Circuit schematic and gain S21 (Meas is measured data, Sim is simulated data) of a 3-stage common source SOI CMOS technology amplifier. An on-chip Through-Reflect-Line (TRL) calibration was done for probe tip measurements.
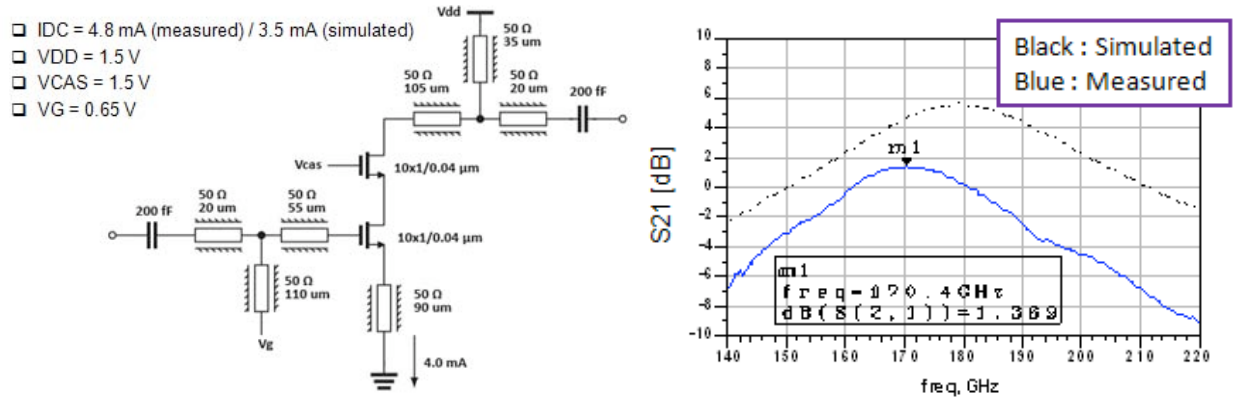
Figure 1 (b). Circuit schematic and gain of a 3-stage common-source SOI CMOS technology amplifier. On-chip TRL calibration was done for probe tip measurements.
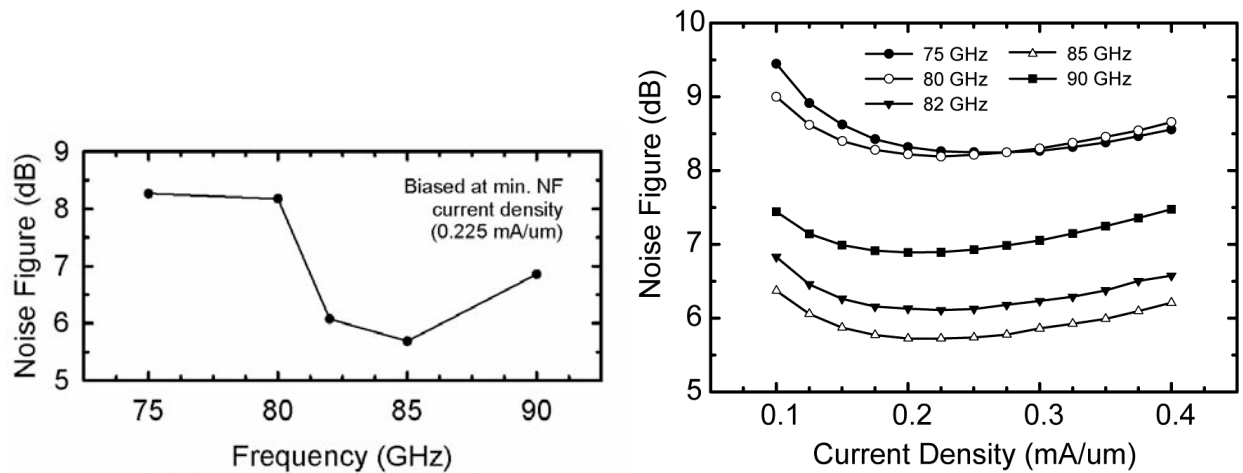
Figure 2. Measured NF of the 3-stage amplifier in Figure 1. A Quinstar W-band noise source and band-limited mixer downconverter and spectrum analyzer was used to measure NF. Measured minimum NF at fixed applied drain current is plotted against frequency, as well as NF versus applied drain current density for different frequencies.
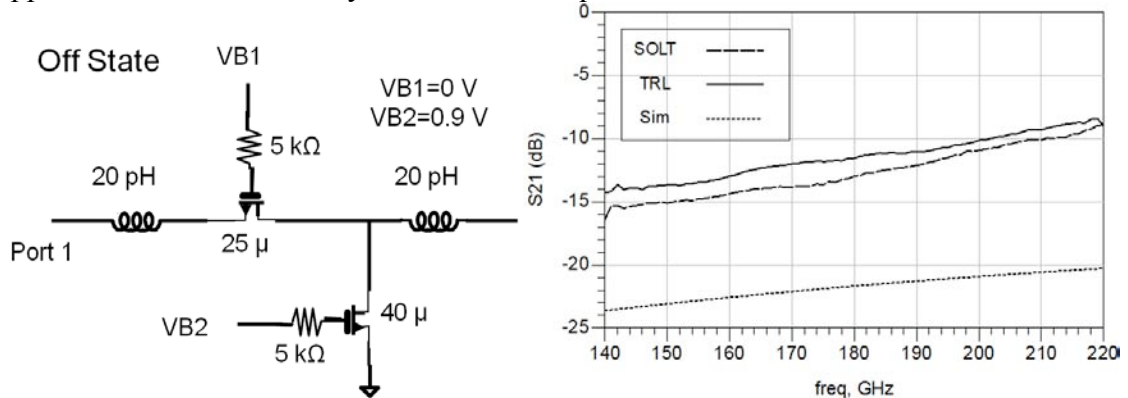
Figure 3 (a). Circuit schematic and plot of insertion loss measurement of the SPST switch in the OFF state (VB1=0V and VB2=0.9V) – isolation. The two calibrations TRL and Short-Open-Load-Thru (SOLT) are measured data. The SOLT calibration is done with an alumina GGB Inc. calibration substrate to set the reference plane at the probe tips. Difference in the two measured data indicate the error and uncertainly in the measurements.
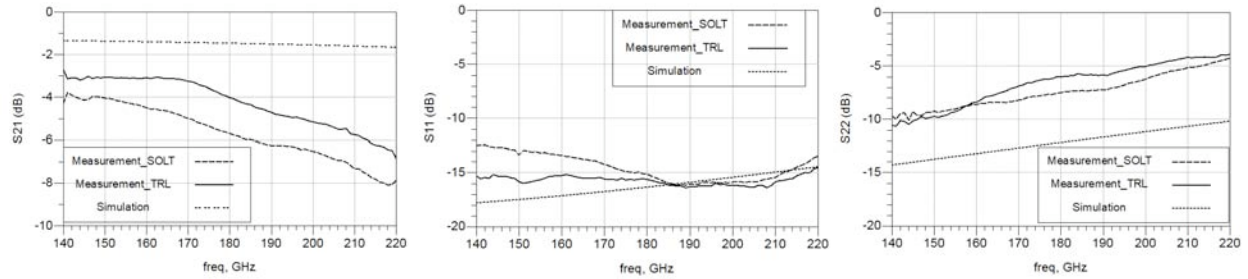
Figure 3 (b). S-parameters of the SPST switch in the ON state (VB1=0.9V and VB2=0V).
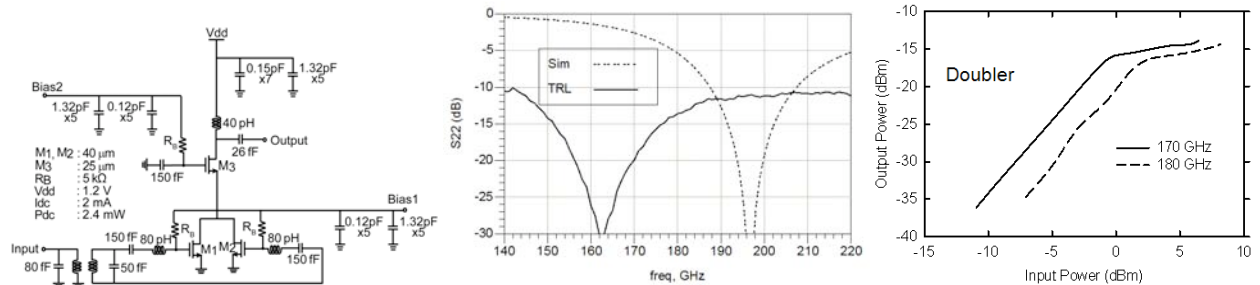
Figure 4. Circuit schematic of a SOI CMOS technology frequency doubler. S-parameter shows the output match (S22) tuned lower in frequency than the simulated design. Measured output power at 170 and 180 GHz are plotted against input power at half the output signal frequencies.
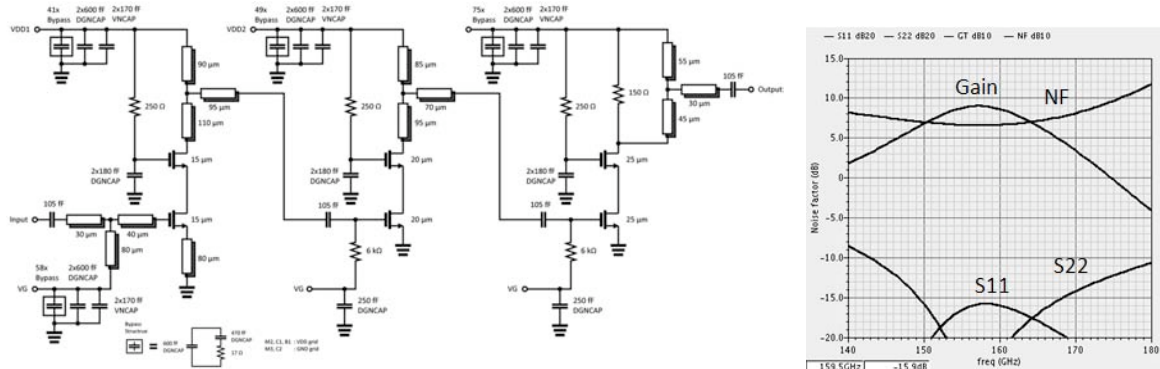
Figure 5(a). Circuit schematic of a second-iteration 160-GHz 3-stage cascade amplifier and its simulated performance characteristics.
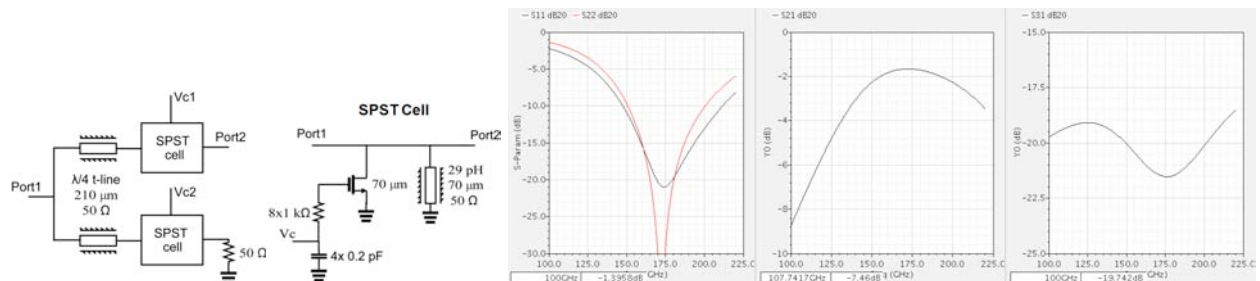
Figure 5(a). Circuit schematic of a second-iteration 180-GHz 3-stage SPDT switch and its simulated performance characteristics.

MICROINSTRUMENTS AND SMALL PAYLOADS

FIRST DEMONSTRATION OF A FLIGHT-LIKE DIGITAL SPECTROMETER FOR FUTURE PLANETARY MICROWAVE SPECTROSCOPY MISSIONS

Robert F. Jarnot (PI), Instrument Systems Implementation and Concepts Section (382)
Mark A. Allen, Planetary Sciences Section (322)
Robert A. Stachnik, Earth Atmospheric Science Section (328)
Dan Werthimer (Co-PI), Space Sciences Laboratory, University of California at Berkeley

A.   OBJECTIVES

The primary objectives of this effort were to design, implement, and demonstrate a complete digital spectrometer subsystem suitable for future planetary microwave spectroscopy missions. Secondary objectives were to demonstrate a spectrometer of this type that provided improvements over competing technologies in terms of bandwidth, resolution, channel shapes, mass, size, power consumption, and manufacturability.

B.   APPROACH AND RESULTS

1.   Approach

The starting point for this effort was to determine requirements, which can be summarized as a digital spectrometer with the following characteristics:

a.   Instantaneous bandwidth of 750 MHz, with resolution of 183 kHz (4096 channels).

b.   Low mass, volume, and power consumption (less than 2 kg, 600 cm^3 and 5 W).

c.   Ability to operate in the radiation environment presented to instruments operating in Mars or Venus orbit for a period of 2 years.

In addition to total dose radiation requirements, we required the spectrometer to be immune to destructive single-event latchup (SEL), and to be designed in a manner such that radiation-induced single-event upsets (SEU) resulted in only temporary data corruption, with the effects of the upset removed (flushed) by normal spectrometer operation.

The key new element in this spectrometer is a custom application-specific integrated circuit (ASIC), which implements the digital signal processing (DSP) functions in the form of a polyphase filterbank [A,B]. Another crucial component is the analog-to-digital

converter (ADC) needed to digitize the analog input signal at a 1500-MHz rate. During this development, a suitable rad-hard, low-power device became available from National Semiconductor (ADC08D1520QML), and was used in this implementation.

A complete prototype spectrometer has been designed and fabricated under this task, but the fabrication activity was unfortunately delayed because delivery of the critical ADC occurred 4 months after its promised due date. This late delivery required us to perform an early evaluation of the ASIC portion of the spectrometer in a hybrid test setup, using an external ADC and FPGA board to provide full spectrometer functionality. A limitation of this test setup was that the board-to-board communications rates limited the sample rate of the test setup to 800 MHz, with a consequent reduction in bandwidth to 400 MHz. Extensive testing of this hybrid system was performed, indicating full functionality of the system, the desired channel shapes, excellent stability, and the ability to perform long measurements in which the noise on spectral data integrates down to the levels needed to support the most stringent science measurements.

Although the prototype flight-like spectrometer will not be completely debugged and characterized until after the end date of this task, sufficient test data have been obtained on the ASIC for us to feel confident that we will be able to demonstrate a full-performance spectrometer within the next 6 months.

2.   Additional Results

Initial testing of the flight-like spectrometer assembly (Figure 1) indicated correct operation when run at full speed (1500-MHz sample rate) when the data source was a built-in test vector generator (Figure 2), but glitches in the data were present when reading actual ADC data (Figure 3). This anomaly is under investigation, and appears to arise due to a difference in wire length between the package pins and the ASIC die for certain data lines. Current efforts to resolve this issue include varying the ASIC supply voltage, altering the ADC data to clock phasing, and adding small capacitances to some data lines to try to match the delivery times of all data bits. If this effort is unsuccessful, an alternative mode of operation of the ASIC that allows us to double the data path width, and halve the transfer rate, has been verified. If this solution is needed, the printed circuit board will require minor redesign, which will be performed outside the scope of this DRDF, but will result in a digital spectrometer meeting the proposed bandwidth requirements.

In addition to functional testing, the ASIC has been tested for SEU and SEL at the Berkeley cyclotron under guidance from the JPL parts group. These tests indicate the absence of SEL, and full automatic recovery from SEU-induced errors. These tests were performed by subjecting the ASIC to various heavy ions while performing end-to-end data processing using the built-in test vector generator (TVG) as the data source. The ASIC includes sophisticated built-in test logic, which allows key nodes to be monitored or forced, and large building blocks (such as the polyphase filters and FFT) to be bypassed. This monitoring ability was used to determine which portions of the ASIC were most susceptible to SEU events, which as expected, were the RAM-rich FFT blocks. These tests also verified that a key design requirement was met

— no individual SEU event was found to corrupt more than a single frame of data before it was automatically cleared.

Mass, volume, and power consumption of the spectrometer module all meet original specifications, and even if it proves necessary to increase the ASIC supply voltage slightly to remove the data glitches, or to redesign the printed circuit board to use four lanes of data from the ADC to the ASIC (instead of the current tow lanes), we expect to still meet our power goals.

3.   Summary

We have developed and demonstrated all of the key functions of a compact, low-cost, low-power digital spectrometer suitable for a range of low cost interplanetary microwave spectroscopy missions.

C.   SIGNIFICANCE OF RESULTS

The technology developed here provides JPL with a competitive advantage when proposing future low-cost interplanetary microwave limb sounders such as VESPER and SOAR due to the significant reduction in mass and power, as well as superior performance (combination of bandwidth and resolution) compared to alternative technologies. In addition, as a successor to this development, we are looking into the possibility of merging the digital spectrometer electronics onto a highly integrated FPGA-based C&DH board to further reduce instrument mass and volume.

Alternatives to the digital spectrometer developed here include acousto-optical (AOS), chirp-transform (CTS), digital autocorrelator (DACS), and analog filterbank spectrometers, all of which have been demonstrated in a flight environment. With the exception of the DACS, none of these alternatives can meet the ensemble of performance requirements listed at the beginning of this report. The significant advantages that this implementation offers over the DACS are in regard to channel shape, signal to noise, and linearity. The channels of a DACS have a sinc function responses with considerable ringing and channel-to-channel overlap, whereas the polyphase spectrometer has user-definable channel shapes which in this implementation have no ringing, deep and steep skirts, and very little channel-to-channel overlap (Figure 4). Unlike the spectrometer developed here, the DACS also suffer from a small signal-to-noise loss (24% to 40% for the units which have been flown to date), and significant inherent nonlinearity that requires correction during data processing.

In addition to its superior performance, the polyphase digital spectrometer is readily manufacturable, and is inherently very stable (due to its digital nature). Since its performance is almost completely defined by digital algorithms, it is largely insensitive to thermal variations, and requires minimal calibration.

D.   NEW TECHNOLOGY

The new technology developed under this task is the low-power digital spectrometer ASIC, which was designed to withstand the radiation environment of VESPER and SOAR types of missions.

E.   FINANCIAL STATUS

The total funding for this task was $200,000, all of which has been expended.

F.   ACKNOWLEDGEMENTS

G.   PUBLICATIONS AND PRESENTATIONS

[A]   Brian Richards, Nicola Nicolici, Henry Chen, Kevin Chao, Dan Werthimer, and Borivoje Nikolić, "A 1.5GS/s 4096-Point Digital Spectrum Analyzer for Space-Borne Applications," presented at the IEEE Custom Integrated Circuits Conference, San Jose, California, September 13–16, 2009.

H.   REFERENCES

[1]   Brian Richards and Prof. Borivoje Nikolić, "The Mars Spectrometer: Soft-Error Tolerant DSP Mapped from Simulink to Silicon," Berkeley Wireless Research Center Winter Retreat, 2009. Download at: http:// bwrc.eecs.berkeley.edu/ php/pubs/pubs.php/899/Brian_Richards_BWRCRetreat2009%20bcr%20v2.pdf

I.   APPENDIX: Field Demonstration of the Digital Spectrometer

In order to demonstrate operation of the spectrometer in the field, the 400-MHz bandwidth preliminary test unit was take to the Goldstone Deep Space Communications Complex and used to observe spectral emission from the Orion Nebula. A plot of the spectral data is shown in Figure 5. As soon as a full-bandwidth prototype is available at JPL, we will conduct further, more challenging, field measurements.

J. FIGURES
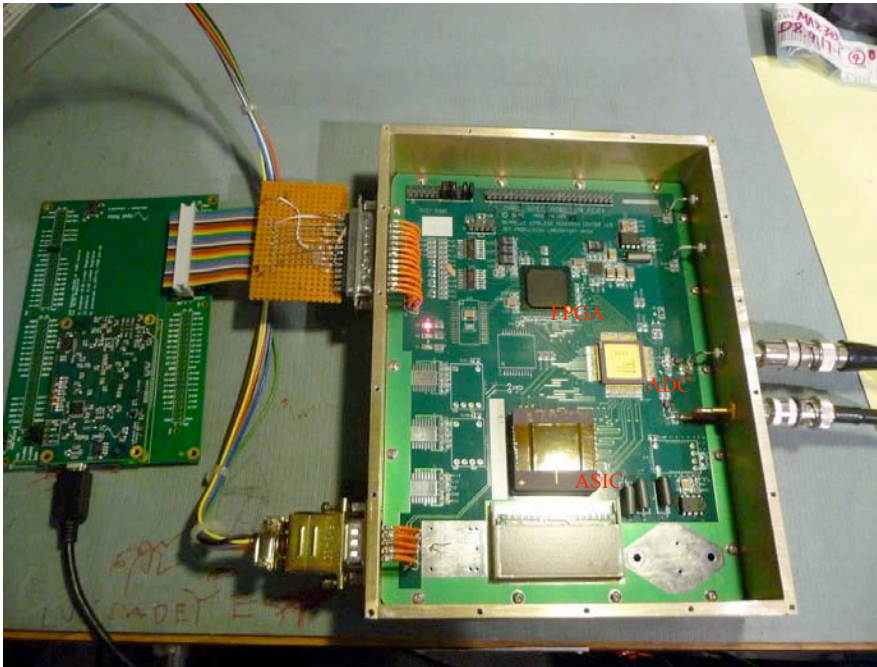
Figure 1. Photo of the prototype flight-like digital spectrometer module. The ASIC and ADC have been labeled. The FPGA provides an interface between the instrument C&DH and the ASIC, ADC, and clock generator (phase-locked loop) in this module.
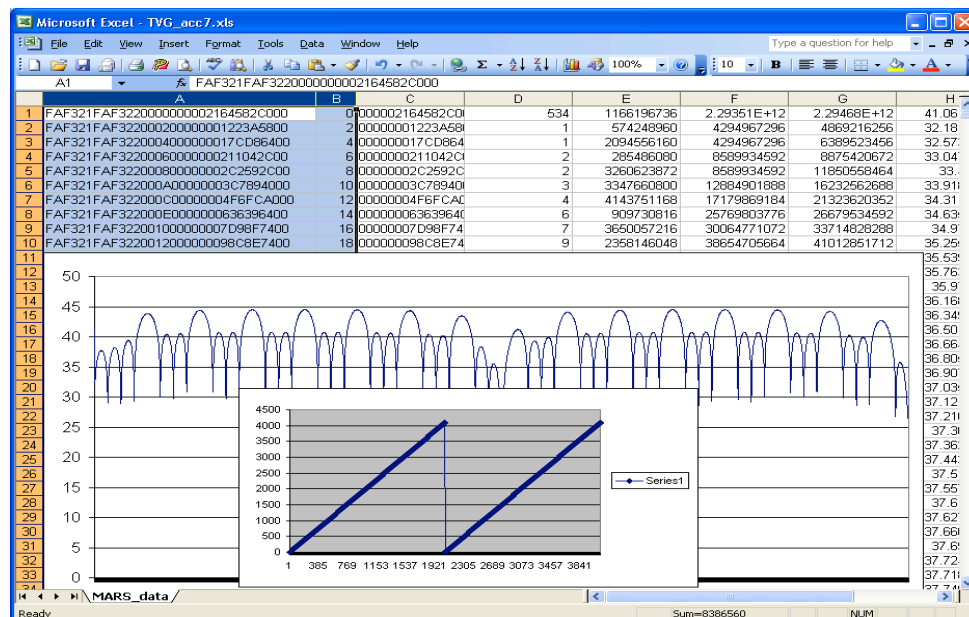
Figure 2. Screen shot showing output data from the spectrometer module when operating at the full 1500-MHz sample rate, but with data provided by an internal test vector generator. The output spectrum is as expected, indicating correct functionality of the entire signal processing chain.
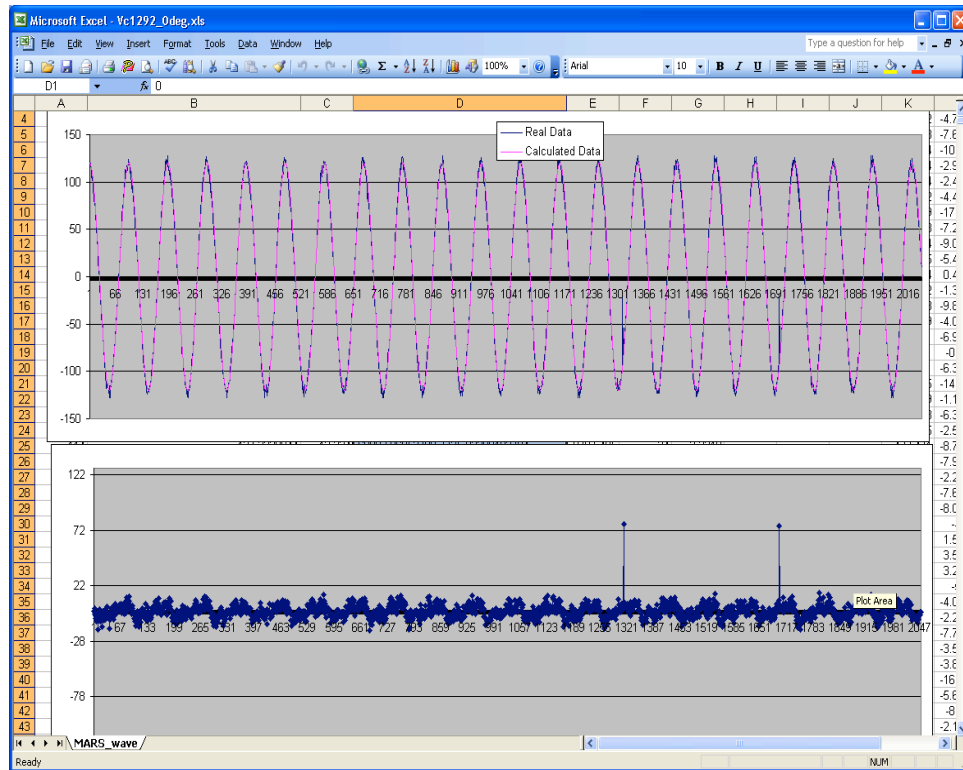
Figure 3. The upper panel shows ADC data sampled by the spectrometer ASIC at 1500 MHz, and the lower panel shows the difference between expected and actual data. The two spikes in the lower panel indicate that bit 6 of the ADC data is occasionally read incorrectly. The remaining ripple in the lower panel reflects the finite accuracy of the ADC.
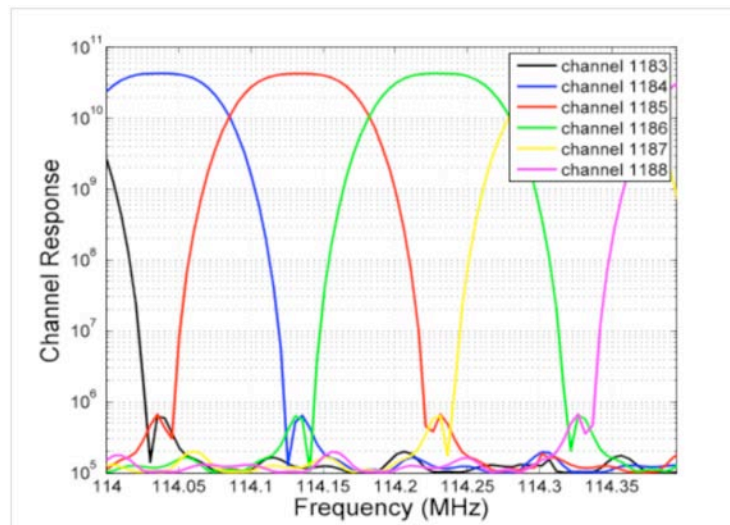
Figure 4. Measured channel shapes of the digital polyphase spectrometer. Note the steep and deep skirts, lack of significant sidelobes (ringing) and small channel-to-channel overlap (the vertical axis has been plotted on a logarithmic scale). Spectral performance of this caliber can only be achieved with a digital polyphase spectrometer of the type developed here.
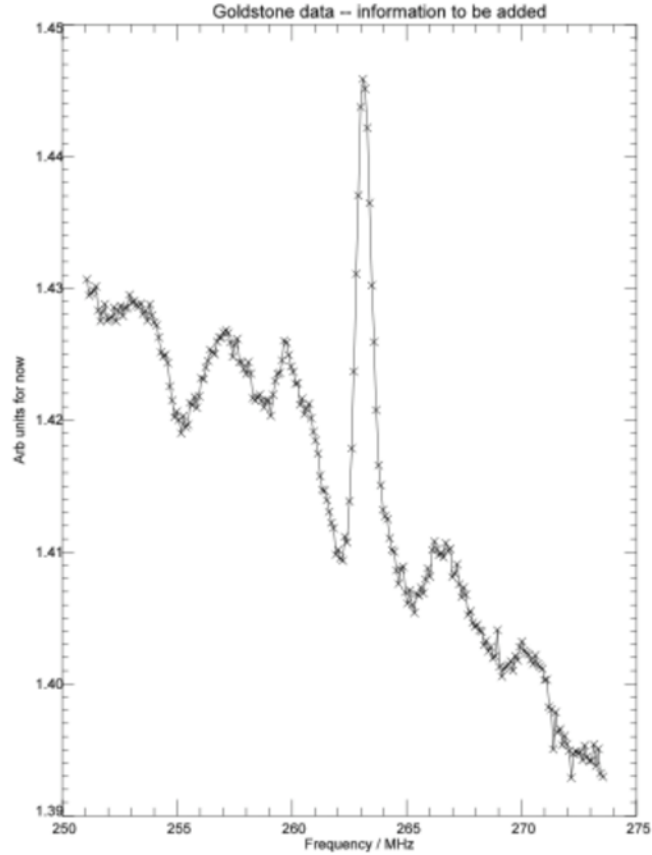
Figure 5. Measurement of an emission line from the Orion Nebula taken at Goldstone using a preliminary version of the digital spectrometer operating with 400-MHz bandwidth. This measurement will be repeated when the full-bandwidth version of the spectrometer is available, and will be made in parallel with the spectrometer routinely used for such measurements for comparison.

SCIENTIFIC CMOS CAMERA FOR COMET AND ASTEROID DETECTION

Stythe T. Elliott (PI), Instrument Electronics and Sensors Section (389)
Larry Hovland, Instrument Electronics and Sensors Section (389)
Stephen S. Murray, Harvard-Smithsonian Center for Astrophysics

A.   OBJECTIVES

The objective of this activity was to demonstrate that CMOS sensors that combine the random access of CMOS imagers with ultra-low-noise detectors are suitable for optical instruments like Whipple, which to date use low-noise charge-coupled device (CCD) detectors. Optical instruments such as Whipple are designed to detect Kuiper Belt, Oort Cloud, and Sedna-like objects through occultations of remote stars. The size of the optics in this type of instrument is determined by the signal-to-noise ratio needed to detect the occultations. The denominator of the SNR is set by the read noise of available detector arrays. These instruments also require the sampling of small windows of interest around several thousand stars within a field of view of millions of pixels and detecting when the stars appear to dim due to a transit of an intervening object.

CCDs have a demonstrated a read noise of several electrons when read out very slowly, but because they read out by sequentially shifting charge from pixel to pixel, they need to read out the whole array to get to the windows of interest. At the sample rates required, this would require the CCD to operate very rapidly, where the noise goes up to 20 electrons or more. The pixels in CMOS detectors can be randomly accessed, but until now their read noise floor was higher than CCDs, in the range of 20 electrons or higher even at slow speeds. Sarnoff Corporation (SAO) has recently developed a scientific CMOS imager with CCD-like read noise. JPL has significant experience with Sarnoff devices. Experiments done at JPL on a prototype device have indicated a noise floor as low as 2 electrons. These experiments suggest that the CMOS detectors made by SAO can produce performance similar to that of a CCD detector array. Taking advantage of the random access capability of the CMOS detectors, we proposed to demonstrate windowing on a number of pre-set windows, while providing the low noise performance at the required pixel readout speed.

B.   APPROACH AND RESULTS

Our goal was to develop a proof-of-concept demonstration of the capabilities of the CMOS detectors made by SAO and the windowing on a number of pre-set windows, while providing the low noise performance at the required pixel readout speed. To achieve this goal, we completed the following steps.

(a)   SAO provided JPL with Sarnoff low-noise CMOS Minimal Imagers (JAZZIVb 5T PPD: Figure 1). Under another task, these devices will be characterized in JPL's current low-noise CCD/CMOS screening camera, including operating parameters such as read noise, dark current, charge transfer efficiency, etc. (see Figure 2).

(b)   JPL designed and fabricated a low-noise (2e-) CMOS-based camera system and ground support electronics (GSE) (see Figure 3, electrical block diagram; and Figure 4, layout of the camera board).

(c)   JPL conducted system testing on the low-noise CMOS camera and validated that all the signals are present and free of noise. JPL further validated that the camera board is capable of receiving lab-view instructions, collecting full-frame data.

The JAZZIVb 5T PPD. detector in the progressive scan readout mode was used to characterize the camera at 50 Kps and 25C (see Figure 5). Windowing was demonstrated through use of a focused pinhole on a 16×8 pixels window. Slides 1 and 3 are a series of the windows. A pen was used to block the beam in front of the pinhole source (note areas where there is no light). Detector functionality was verified by moving the pen. Figures 6 and 7 are column traces taken through the point source while on and off.

At SAO, we designed and fabricated an optical test system to simulate occultation events produced by Kuiper Belt objects, Sedna-like sources, or Oort Cloud objects. The optical simulator takes as input an idealized light curve for a particular event and generates an LED light signal that corresponds to the input light curve in time and amplitude. The LED star image is projected onto the optical detector using fiber optics and lenses. The resulting star image simulates the expected image from the Whipple telescope. We also designed and build the address sequencer for running the Sarnoff CMOS minimal array, and the data interface from the JPL CMOS camera to the data collection computer. Event recognition algorithms have been fully implemented and tested using simulations derived from the idealized occultation light curves.

Figure 8 shows an idealized light curve, the simulated output from a CMOS minimal due to this input light curve, and the output of the event recognition algorithm. We will compare these simulated results with the actual CMOS cameral performance and also analyze the detected event to recover the key parameters of object size and distance.

SAO provided test CMOS minimal arrays (via Sarnoff) to JPL for the camera development.

The JPL CMOS camera is not complete, and so there has been no end-to-end testing of the camera, data system, and event simulator.

C.   SIGNIFICANCE OF RESULTS

The successful demonstration of low-noise CMOS detectors would allow the development of a new generation of optical instruments (such as Whipple-II) with a substantially

reduced cost due to the reduced size of the collecting mirror enabled by the superior performance of these detectors.

D.   FINANCIAL STATUS

E.   ACKNOWLEDGEMENTS

F.   PUBLICATIONS

None.

G.   REFERENCES

None.

H.   FIGURES

Figure 1. Cross-section of a SAO CMOS detector**.**

Figure 2. Optical response of the SAO CMOS detector.
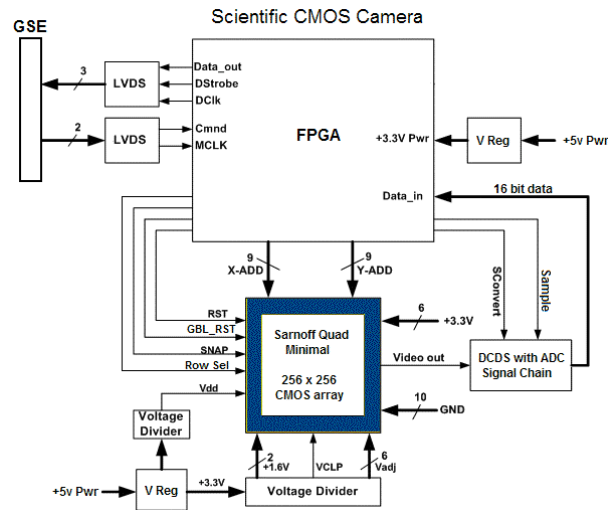
Figure 3. Block diagram of the CMOS camera board.
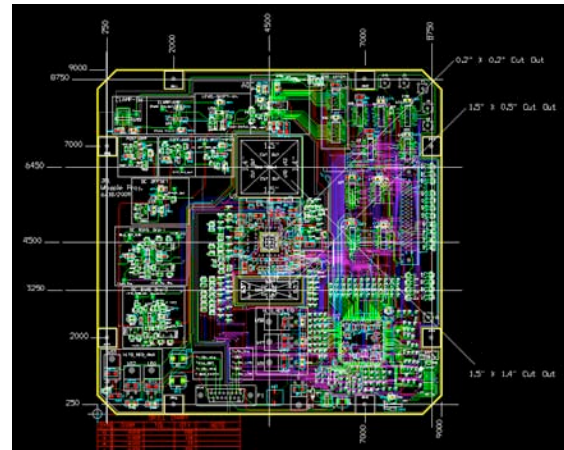
Figure 4. CMOS Camera Board layout.
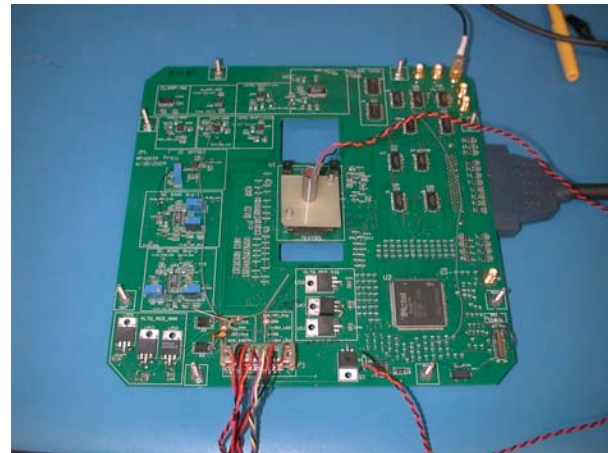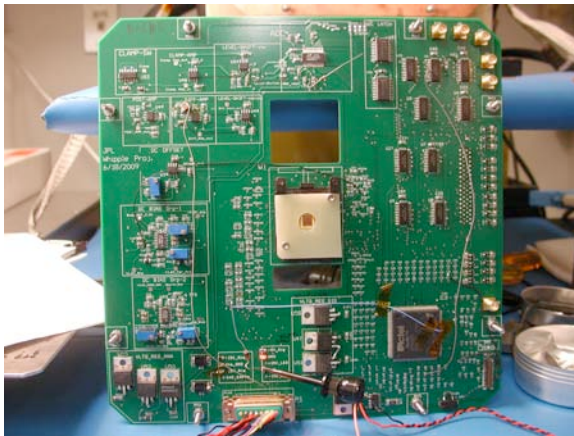
Figure 5. CMOS camera board with SAO detector and light source.
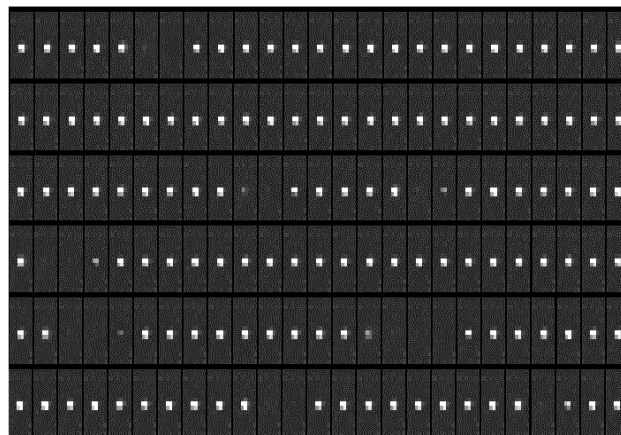
Figure 6. Point source located in 16×8 window showing occultations.
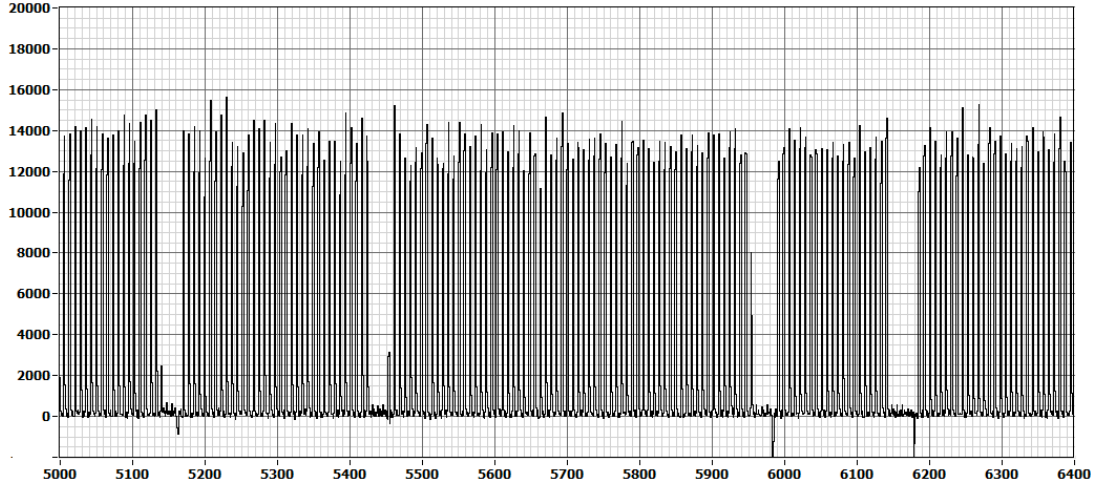
Figure 7. Column trace through windowed area showing point source going off and on as beam is blocked.
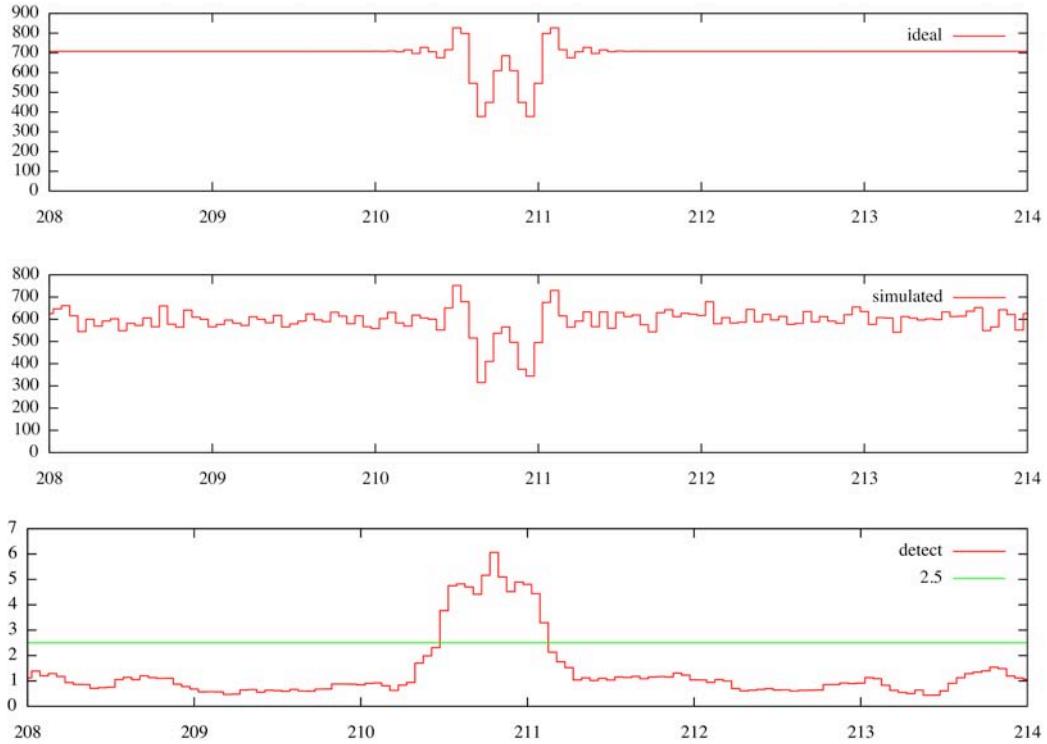
Figure 8. The upper panel is an idealized light curve for a Sedna-like source with a 5-km diameter at 500 AU distance occulting a magnitude 14 star. The sampling is at 20 Hz. The middle panel is the simulated light curve as detected by the CMOS minimal array, including shot noise, pixel read noise, pixel-to-pixel variations, etc. The bottom panel is the output from the event detection algorithm (equivalent width), where the green horizontal line is the threshold for triggering a detected event.

DEVELOPMENT AND TESTING OF A PROTOTYPE
OUTER PLANETS MAGNETOMETER

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1330

Carol A. Raymond (PI), Solar System Mission Formulation Office (410)
Neil Murphy, Astrophysics and Space Sciences Section (326)
Vassilis Angelopoulos, Institute of Geophysics and Planetary Physics,
University of California at Los Angeles

A.   OBJECTIVES

The objective was to develop and test a prototype Vector Helium Magnetometer (VHM) system designed for an Outer Planets Mission, using a compact laser-pumped sensor developed at JPL and an innovative second-order sigma-delta ($\Sigma-\Delta$) signal acquisition technique, developed at UCLA.

A helium magnetometer senses magnetic fields by measuring the change in absorption of light as it passes thorough a glass cell containing helium. The helium is excited by an RF discharge that produces a population of helium atoms in the 2^3S_1 metastable state [1]. Figure 1 shows the relevant energy levels in this process: when circularly polarized light at a wavelength of 1083 nm is absorbed by helium atoms in the 2^3S_1 state in the presence of a magnetic field, only one of the m_J levels can be excited. As absorption and spontaneous emission continue, this leads to an imbalance in the populations of the different m_J levels, which is referred to as optical pumping. The efficiency of this optical pumping depends on the angle between the magnetic field and the direction of light passing through the helium. A vector helium magnetometer operates by applying a rotating magnetic field, via a set of orthogonal coils surrounding the sensor (Figures 2 and 3), which modulates the absorption of light. Any nonzero ambient field at the sensor results in a modulation frequency component at the rotation frequency. This component is used as an error signal in a servo control loop to feed back a current to the sensor coils, nulling the field seen by the sensor. This nulling feedback current, properly calibrated, provides a measurement of the ambient field at the sensor.

By replacing the helium lamp traditionally used as a light source in the VHM with a diode laser, the intrinsic sensitivity of the magnetometer can be increased significantly. This also allows us to produce a much smaller sensor, with sensitivity equivalent to current flight requirements, which is the goal of this effort. Combining this sensor with innovative electronics for signal acquisition, instrument control, and signal processing will produce a low-mass, low-power magnetometer, ideally suited for use in mass-constrained missions.

B.   APPROACH AND RESULTS

     1.   System Description

       The system components designed, developed, and tested as part of this activity were: a) a compact laser-pumped VHM sensor, b) a control electronics board, and c) a 27-MHz radio frequency (RF) generator.

       a.   The compact VHM sensor was designed to minimize mass and also produce a sensor with a minimal number of parts to make assembly straightforward (a CAD model of the sensor is shown in Figure 2). This sensor was developed from our original laser-pumped VHM sensor, developed as a proof-of-concept instrument, which weighed approximately 200 g and was approximately 4× the volume of the current sensor design. The final sensor fits within a volume of 60 mm × 25 mm × 25 mm and weighs approximately 60 g, not including the mass of cables. Figure 3 shows the assembled sensor, prior to testing.

       b.   A combined analog/digital electronics board was designed and fabricated as part of the DRDF development (Figure 4). The board carried the analog electronics to provide the sweep magnetic field and nulling magnetic field to the sensor coils, a preamplifier for the sensor detector signal, and a floating-point gate array (FPGA) that carried the firmware to operate the magnetometer.

       c.   Several RF designs were fabricated and tested for efficiency and reliability of igniting the required discharge in the sensor cell. Figure 5 shows two of the designs used: on the left a board based on heritage designs; on the right a more efficient design that also uses components that are more radiation tolerant. These boards were used primarily to study the issues that affect the efficiency of the VHM RF system, to carry out measurements of helium cell impedance under various conditions, and to optimize the cable to cell RF matching.

     2.   Testing

       Tests were carried out on the various system components, followed by system tests on the combined sensor and analog/digital electronics board. To support early development, the control electronics were tested using a conventional lamp-pumped sensor and the compact VHM sensor was tested using engineering model electronics from the Cassini mission.

       System-level testing was carried out at JPL after integrating the electronics board and sensor. The sensor was located in a magnetically shielded room, with a large Helmholtz coil system for nulling out any residual magnetic field and applying known fields for test and calibration purposes. This allowed the instrument to be tested over a range of magnetic fields typical of what might be expected in Jupiter's magnetosphere and in a low-noise environment to allow sensor noise characterization. The electronics were coupled to the sensor using a cable assembly, similar to one that would be used on a flight system to couple a boom-mounted sensor to instrument electronics housed within the spacecraft body.

The main goal of the testing program was to demonstrate that the compact sensor design and Σ−Δelectronics would function with a noise level (5 pT/√Hz @ 1 Hz) similar to heritage flight VHMs and also the first-generation (larger) laser-pumped VHM. Figure 6 shows a power-spectral density (PSD) plot of data from one axis of the VHM taken during system testing. The horizontal line is at 4 pT/√Hz and indicates that the system can perform at the expected noise level. The data shown are for the quietest axis, perpendicular to the sensor optical axis – data for the field component parallel to the optical axis have a higher noise level.

Tests of the RF system were carried out independently of the magnetometer system tests and suggest that a radiation-tolerant RF system for the VHM can be produced that uses approximately half the power of current flight VHMs. The RF design is less efficient than previous flight systems; however, there is a significant saving due to the replacement of the helium lamp with the laser diode.

C.   SIGNIFICANCE OF RESULTS

The research undertaken as part of this DRDF task has shown that a compact laser-pumped VHM system can operate at the required sensitivity. The small sensor volume and mass will enable us to produce a low-mass, radiation-hardened sensor for use in the Jupiter system. All the components of the sensor can be manufactured from radiation-tolerant materials and the sensor geometry lends itself to efficient shielding to reduce the sensor radiation dose if necessary. While focused on the requirements of a magnetometer for use in Jupiter's high-radiation environment, the resulting system can be used on any mission that requires magnetic field measurements; this includes missions in heliophysics, Earth science, and planetary science.

D.   NEW TECHNOLOGY

The compact VHM sensor is a significant improvement on previous sensor designs and will enable lower-mass systems to be flown on future missions. Based on this DRDF development, a New Technology Report (NTR #47665) has been filed.

E.   FINANCIAL STATUS

The total funding for this task was $163,800, all of which has been expended.

F.   ACKNOWLEDGEMENTS

David Pierce of UCLA carried out the electronics design for our DRDF task. David Pierce and Kathryn Rowe, of UCLA, supported tests on the magnetometer system at JPL, together with undergraduate research students Stephanie Zajac and Ivan Bernal. Stephanie Zajac was responsible for producing the LabView data acquisition system used in preliminary testing.

G.   PUBLICATIONS

None.

H.   REFERENCES

[1]   Edward J. Smith, Benjamin V. Connor, and George T. Foster, Jr., "Measuring the Magnetic Fields of Jupiter and the Outer Solar System," *IEEE Transactions on Magnetics, Vol. MAG-11* **4** (July 1975): pp. 962–980.

I.   FIGURES
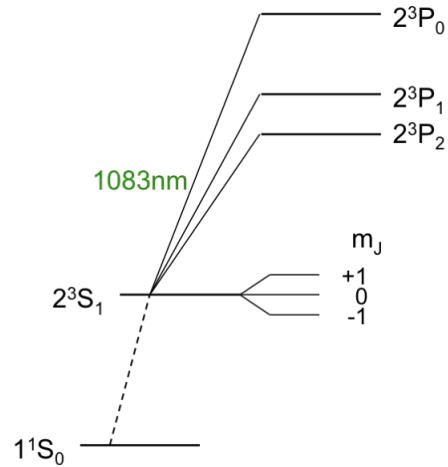
Figure 1. The relevant energy levels of the helium atom (not to scale). Optical pumping takes place when circularly polarized light at a wavelength of 1083 nm is absorbed by atoms in the 2^3S_1 metastable state.
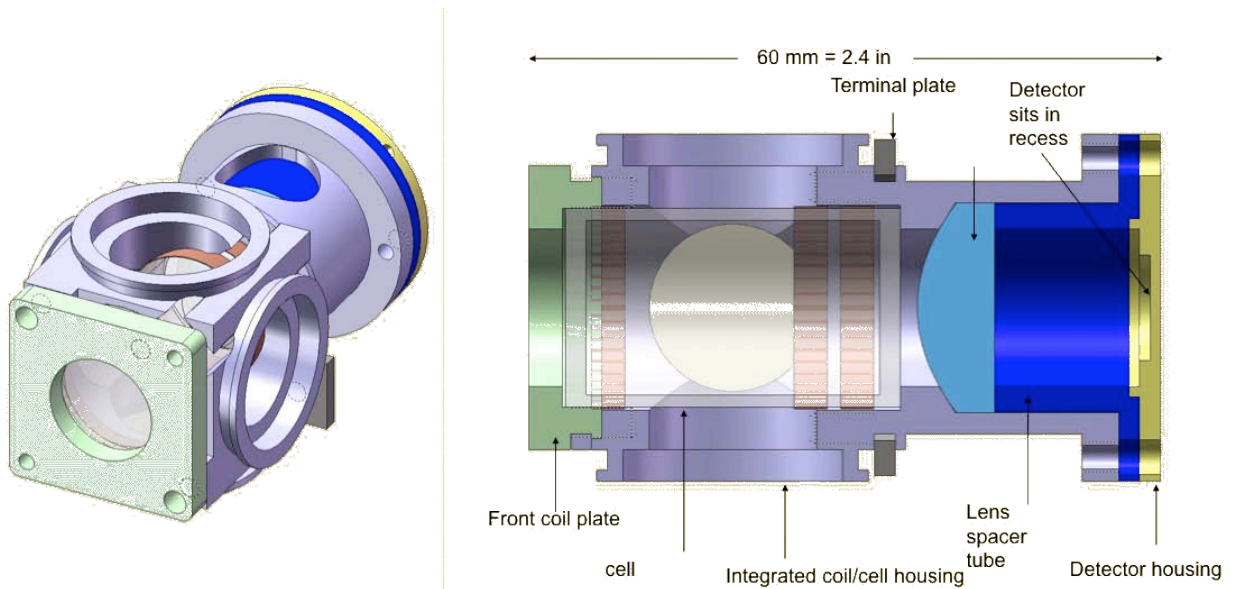
Figure 2. CAD model of the compact VHM sensor.
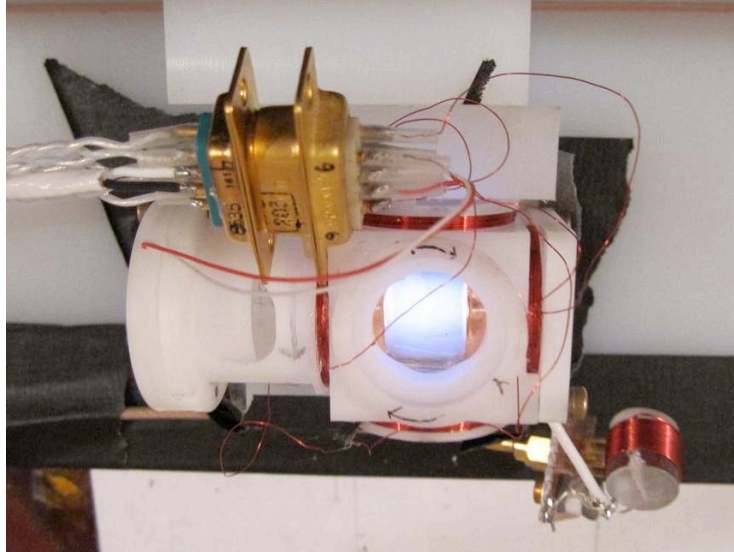
Figure 3. Assembled compact VHM sensor during testing.
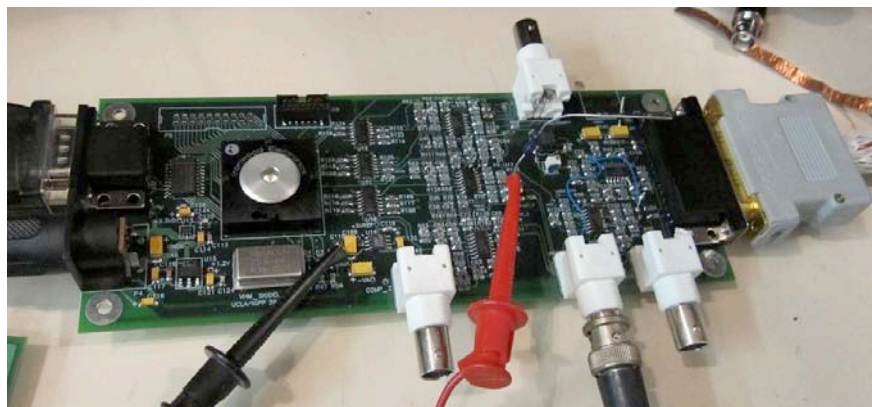
Figure 4. Breadboard of the compact VHM analog/digital electronics. Most functions (including the Σ−Δanalog-to-digital convertor) are implemented in FPGA firmware.
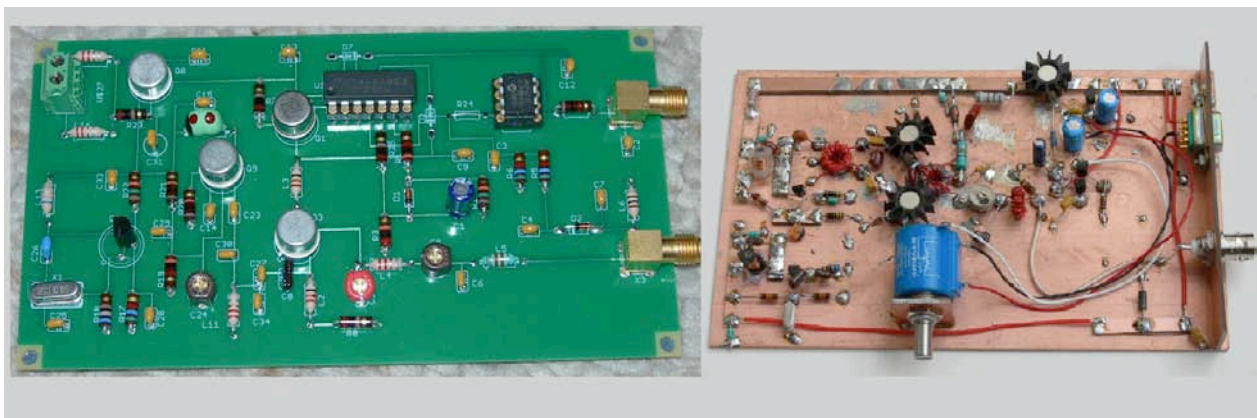
Figure 5. Two of the RF generators used in testing. On the left is a board closely based on heritage designs; on the right is a more efficient design in breadboard form.
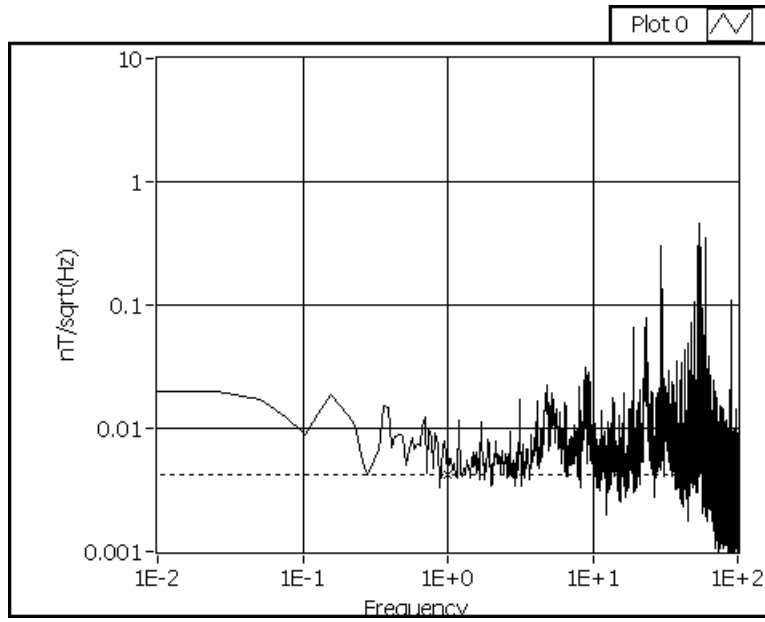
Figure 6. Noise measurements for the compact VHM, taken during system testing. The horizontal dotted line is at 4pT/√Hz.

MISSION DESIGN

GALCIT AEROSPACE RESEARCH AND TEACHING PROGRAM

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1299

Virendra Sarohia (PI), Office of the Chief Technologist (130)
Ares Rosakis, Engineering and Applied Science Division,
California Institute of Technology
Sergio Pellegrino, Aeronautics and Civil Engineering Department,
California Institute of Technology
Olive Stohlman, GALCIT

A.   OBJECTIVES

The main objective of this effort was to engage aerospace engineering graduate students in the fundamental field of fluid mechanics, solid mechanics, material structures, and propulsion with emphasis in modern applications in space technology of strategic importance to JPL. The area of emphasis included, but was not limited to, large lightweight structures, electrical propulsion, spacecraft navigation, spacecraft design, system engineering, and entry/descent and landing, as examples. Under supervision of JPL instructors and Caltech faculty, graduate students undertook research on elastic and thermal performance of large deployable space structures related to the Nuclear Spectroscopic Telescope Array (NuSTAR) mission.

This effort was carried out under a three-term graduate course (AE105), a major new teaching initiative at the Graduate Aerospace Laboratories of the California Institute of Technology (GALCIT). The AE105 course description can be found at the Caltech website.

B.   APPROACH AND RESULTS

The course philosophy was to balance the technical breadth necessary in today's engineering projects and the technical depth consistent with California Institute of Technology culture and MS-level analysis expectations, and avoid teaching "system engineering" as subject matter, but rather use "system interactions" as the driver for analysis of key technical issues.

The course used mission phases as the architectural framework for the three-term sequence to identify issues and keep student interest. These phases were (a) launch/cruise phase, (b) flyby/orbital-observation phase, and (c) entry/descent/landing phase. Within each term, the course focused on (a) methods for identifying driving requirements and key issues; (b) methods of analysis, including analytical and numerical, standard packages, etc; (c) matching fidelity of analysis with phase in project lifecycle; and (d) mixing guidance, navigation, and control (GNC), mechanical, and software together in each term, but with examples that emphasize different specialties per term. Homework problem sets were developed to build upon each other

throughout the first two terms and lead to a more complex and challenging project in the third term (Ae105c). We also utilized and required a standard software package such as MATLAB.

The third term of the class had the structure of a JPL Flight Project. A Preliminary Design Review (PDR) was held after four weeks, and a Critical Design Review (CDR) at the end of the eight-week term. Members of the Review Board were Dr. Greg Davis, Dr. Yunjin Kim, Prof. Sergio Pellegrino, Dr. Jay Polk, and Dr. Marco Quadrelli. Distinguished guests included Dr. Jean-Lou Chameau, Prof. Paul Dimotakis, Prof. Guruswami Ravichandran, Prof. Ares Rosakis, and Gen. Eugene Tattini (USAF Ret.).

The final term project in Aerospace Engineering was held during the Spring Term 2009 (Ae105c, a course listed as mandatory for MS in Aerospace Engineering at the California Institute of Technology). Due to its magnitude and the diversity in subsystems to be considered, and for an adequate systems engineering analysis, the class was split into four teams with different tasks: experimental, model identification, structural dynamics, and dynamics.

AE105 class students were divided into four groups: systems, structural, dynamics, and experimental teams. Each team carried out details of analytical ingestions and their results were compared with the experimental data. Teams also identified what future work needs be undertaken to further advance the deployment of the boom under zero gravity.

C.   SIGNIFICANCE OF RESULTS

The AE105 course has proven to be a success, since the lecturers managed to lead the class into producing a complete and integrated product both from a technical and a systems engineering perspective. Key accomplishments of this year's effort are as follows:

Loads: Gravitational, orbital perturbation and reaction wheel disturbance loads were calculated.
Experimental: Compared static and dynamic tests and compared the data with predicted results.
Systems: Used two independent algorithms to successfully analyze the experimental data.
Structural Dynamics: Abaqus model yielded results in agreement with the physical model. Improvements in model expected to capture nonlinearities present in the physical model.

In summary, the AE105 students produced an integrated system model that predicted the response of a NuSTAR-like spacecraft in an orbital environment.

For further details, see the Appendix below.

D.   FINANCIAL STATUS

The total funding for this task was $95,000, all of which has been expended.

E.   PUBLICATIONS

None.

F.   REFERENCES

None

G.   APPENDIX: REPORTS FROM THE STUDENT TEAMS

The student teams characterized an ATK coilable mast, which is a slender, stowable space structure similar to the deployable mast that will be used on the Caltech–JPL NuSTAR mission. The four teams were able to synthesize the experimental and analytical results to produce an integrated system model that predicts the response of a NuSTAR-like spacecraft in an orbital environment. This verified, integrated model has the capability to simulate spacecraft structural response under a variety of scenarios, including pointing control studies and thermal-structural response.

Report from the experimental team (Jason Cerundolo, Pelayo Dominguez, Vivek Viswanathan)

The goals of the experimental team were to physically characterize the geometry and materials of the ATK coilable mast in the laboratory and to identify the lowest bending and torsional modes. This data was delivered to the model identification and structural dynamics teams.

Force-response curves of the mast were produced with an experimental setup including a shaker, load cell, and rangefinding lasers. The canister, which supports one end of the mast, was firmly mounted on a steel supporting structure with ratcheted tie-down straps so that the mast was horizontal to the lab floor. Laser displacement sensors were used to measure vertical and horizontal displacement. A shaker was attached to the center of the mast endplate, and applied vertical lateral loads to the mast tip. The shaker was driven by a signal generator providing a sine-sweep from 5 to 100 Hz. A load cell mounted at the interface between the shaker and the mast measured the force applied by the shaker. The experimental setup can be seen in Figure 1. Figure 2 shows a typical frequency response data to a sine sweep input.

Using the National Instruments LabView software to record the data at 1 kHz, the experimental team was able to identify peaks in the force-displacement curve near 11 Hz and 14 Hz, corresponding to beam bending modes, and a resonance at 52 Hz corresponding to a torsional mode. To collect data for the damping properties of the mast, the team performed a tap test with laser displacement data. This test was later used by the model identification team to identify the damping coefficient of the fundamental bending mode.

Report from the model identification team (Jason Rabinovich, Chi Wan Ko)

The model identification team's primary goal was the advanced analysis of the experimental data. This team analyzed the tap tests to find the damping ratios of the two bending modes and the torsional mode. They also produced estimates of damping through the frequency sweep data, both through the frequency response function and the eigenvalue realization algorithm. This analysis, described in Figure 3, led to the conclusion that damping was nonuniform in the structure.

Report from the structural dynamics team (Julia Cossé, Zack Rubin, Daniel Talancon)

The goal of the structural dynamics team was to build a finite-element model of the ATK mast. Using the Abaqus finite-element modeling package, the structural dynamics team produced a model of six bays of the mast, shown in Figure 4. The model treats all the members as 1-dimensional truss elements, without joint stress and with idealized joints. The material properties for the Abaqus model were measured by the experimental team, and the damping factor was taken from the work of the model identification team. Some unmeasured properties were assumed to fall within the range for fiberglass rods. The Abaqus model reported a resonant bending mode at 14 Hz and a torsional mode at 52.2 Hz. The Abaqus model also produced an axial mode that was used by the dynamics team.

With the addition of damping ratio data from the model identification team, a sine-sweep test could be run on the Abaqus model. The sine sweep input to Abaqus mimics the sine sweep testing performed on the ATK mast by the experimental team, and the results are shown in Figure 5. Both the experimental and simulated tests provide evidence of mode splitting near the first bending mode.

Report from the dynamics team (Alejandro Lopez Ortega, Keith Patterson, Iliya Zilberter)

The dynamics team integrated a simplified model of the mast with an orbital model and a theoretical attitude control algorithm. In their simulation, they assumed that the boom could be represented by a linear, flexible, homogeneous and isotropic beam. The structural properties were set to match the first bending, twist, and axial modes determined by the experimental data and the Abaqus model. Forces were applied in the simulation at the tip of the beam, corresponding to the masses of the optics bench of NuSTAR and the NuSTAR spacecraft bus. A reaction wheel was modeled as input. The rotation speed of the wheel was set to a constant 4000 rpm, based on a common regime for these elements of 3000 to 4000 rpm.

The dynamics team modeled the mast as it would fly. The mast used in the laboratory experiments was a six-bay model of slightly under 1 meter length, while space telescope applications may call for 10-meter-long masts. Further, the space-based model calls for a free-free boundary condition for the mast, with added point masses at either end. The laboratory mast had a clamped-free boundary condition. In order to translate between these two scales and boundary conditions, simple beam theory was used to extrapolate the laboratory and Abaqus results to a full-scale structure. The results of one dynamic simulation are shown in Figure 6. The

work of the dynamics team was the final step in the path from characterizing a physical mast to predicting on-orbit behavior.

H. FIGURES
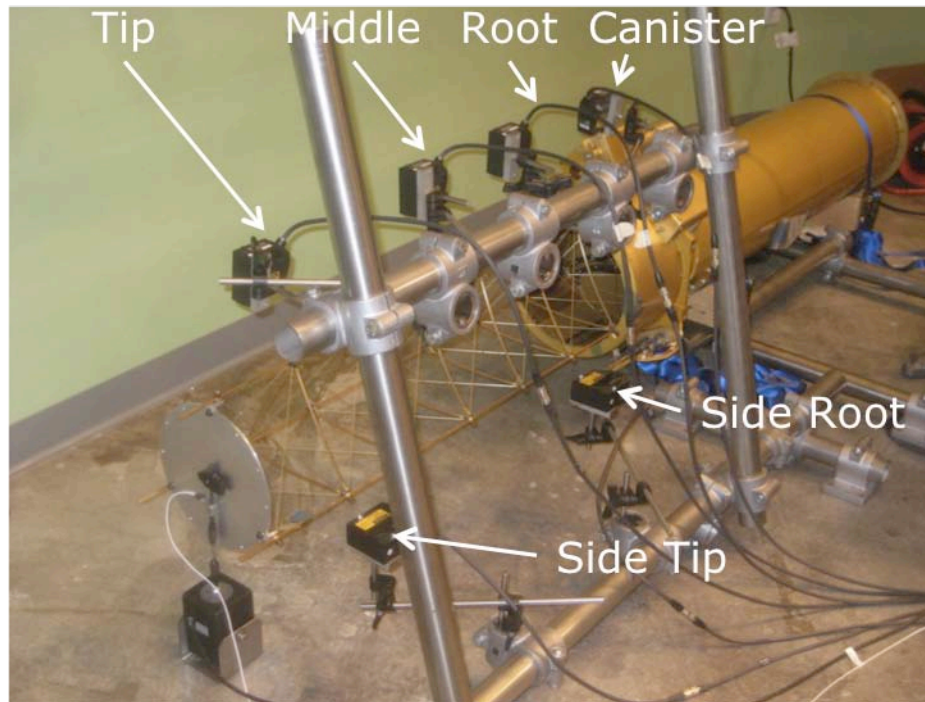
Figure 1. Experimental setup.
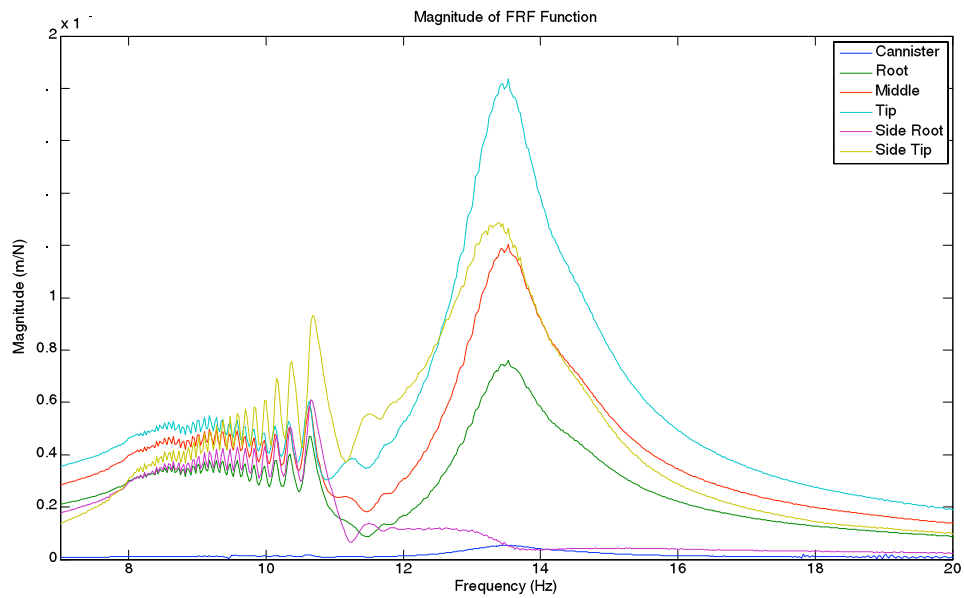
Figure 2. Frequency response function at six points on the boom (experimental).
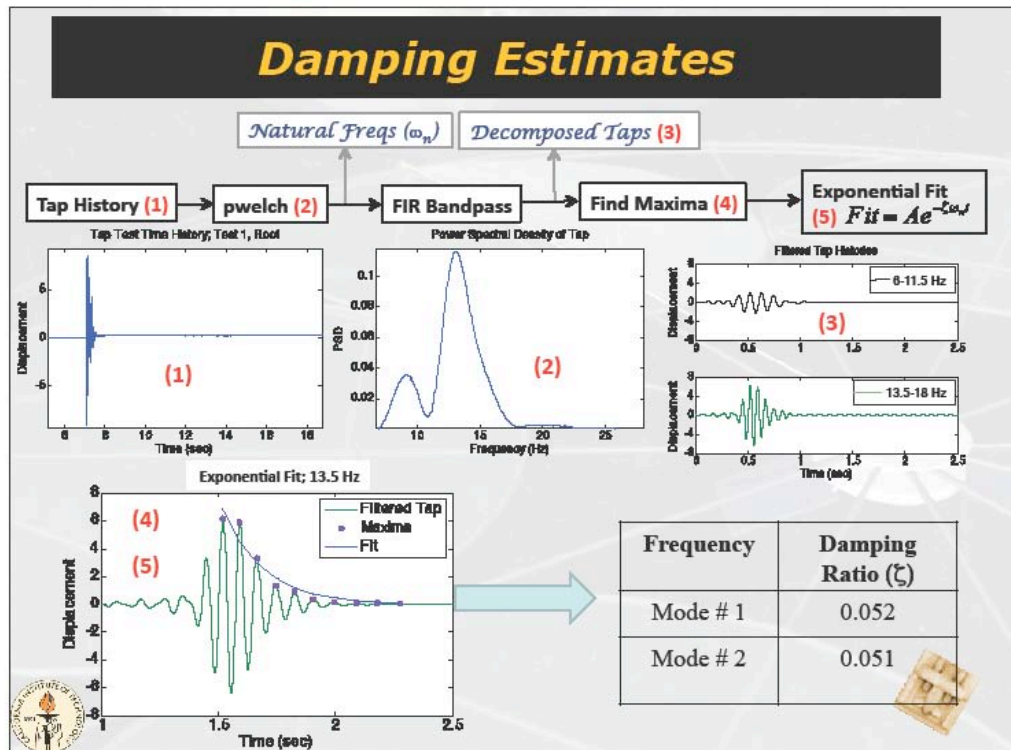
Figure 3. Damping ratio estimation methodology adopted by model identification team.
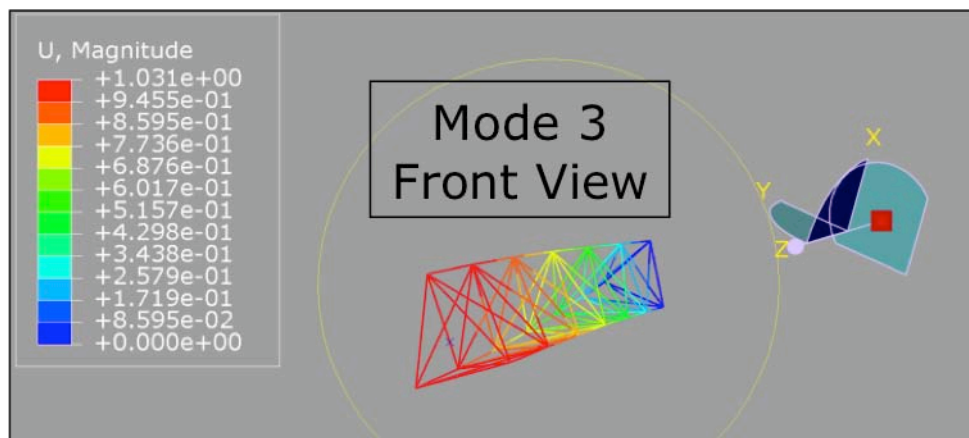
Figure 4. Abaqus model developed by dynamic analysis team.
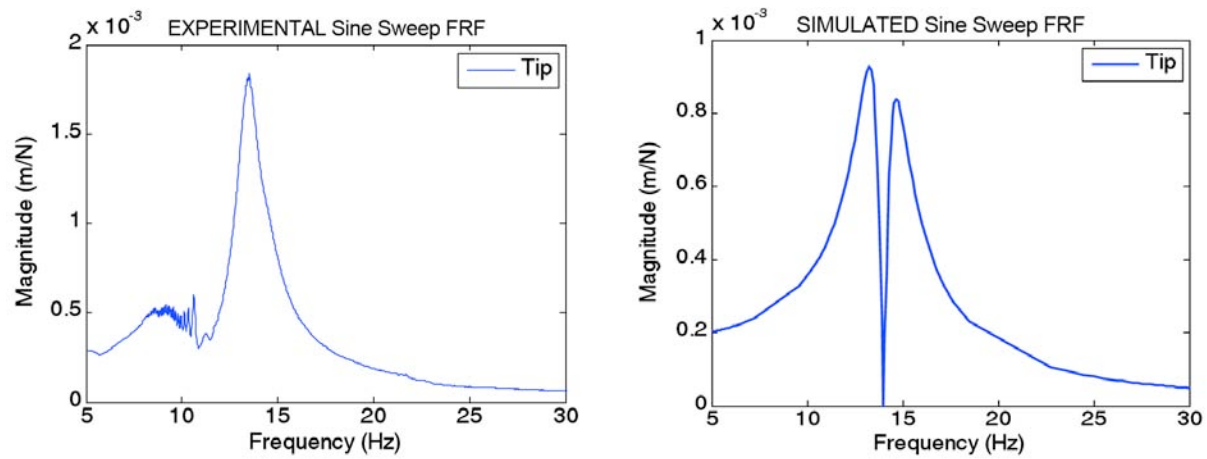
Figure 5. Experimental and simulation frequency response functions to sine sweep input.
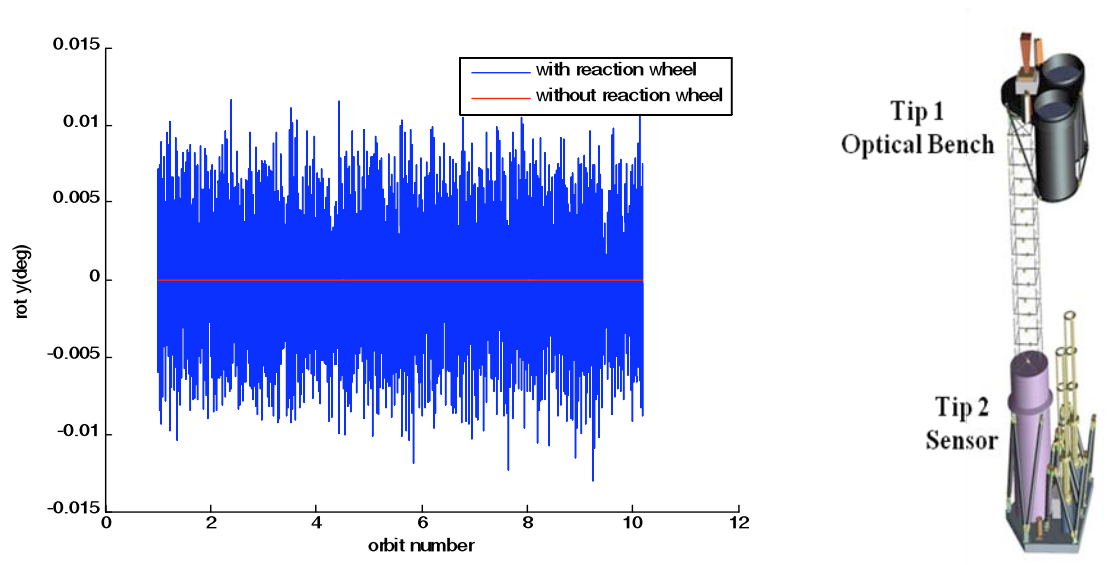
Figure 6. Dynamics team result for pointing errors caused by reaction wheel.

PROPULSION

HIGH-ENERGY IONS IN THE DISCHARGE CHAMBER OF ION ENGINES

James E. Polk (PI), Propulsion and Materials Engineering Section (353)
JoHanna N. Przybylowski-Silva, Graduate Aerospace Laboratories,
California Institute of Technology
Joseph E. Shepherd, Graduate Aerospace Laboratories, California Institute of Technology

A.   OBJECTIVES

The objective was to investigate the ion and plasma properties in a model discharge chamber of an ion engine so that correlations between the ion properties and the plasma properties could be identified. In order to facilitate this goal, a secondary objective was to design an instrument to filter ions based both on their velocity and on their energy-per-charge.

B.   APPROACH AND RESULTS

Task 1: Observe high-energy ions at discharge conditions that are representative of the operating parameters of ion engines.

Using an electrostatic energy analyzer, an instrument that filters ions based on the ratio of their energy to charge, ions with energy-to-charge ratios nearly 10 times that of the discharge voltage were observed. On a range of transmission energy-to-charge ratios from 0 V to 200 V, there are three distinct ion populations and a high-energy ion tail. The high-energy ions are more prominent when there is a lower xenon flow rate into the cathode. The ion energy distribution at the discharge current of 16 A and cathode flow rate of 3 sccm is particularly interesting due to three distinctive peaks in addition to the high-energy tail. As shown in Figure 1, the peaks occur at 24 V, 35 V and 85 V.

Task 2: Design an instrument capable of simultaneously filtering ions based upon their velocity and filtering ions based upon their ratio of energy-to-charge.

The velocity filter consists of perpendicular magnetic and electric fields that allow only ions with a certain velocity to pass through. It was designed using an iterative approach that (1) modeled the magnetic and electric fields given a prescribed geometric arrangement of the magnets, electric plates, and structural materials, and which (2) simulated ion trajectories through these fields. Computer simulations indicate that on the operating range of 1200–4000 m/s, the resolution is 15%.

The energy-to-charge filter, often referred to as an electrostatic energy analyzer, consists of two concentric hemispheres having different applied voltages that create a radial electric field. Ions entering with sufficient energy such that the centrifugal acceleration balances the electrostatic force will pass through, reaching the detector.

As of 9/23/10, only data with the energy-to-charge filter were gathered; an electron multiplier was appended to this filter in order to amplify the low current signal. Machining of the velocity filter components was nearly complete.

Task 3: Construct a set of probes to monitor the plasma properties.

Two shielded Langmuir probes were placed inside a modified conical anode. An inner tungsten wire was surrounded by a ceramic tube, which in turn was surrounded by a steel tube and an outermost ceramic tube. The shielding was added to reduce the signal noise introduced by the instabilities in the discharge plasma through which the probe must pass. One probe was mounted on two perpendicular stages, which enabled it to be translated along the cathode axis and perpendicular to the axis.

Comparing the signals gathered simultaneously on the stationary probe and on the moving probe enabled an investigation into the nature of the waves in the discharge plasma. Figure 2 displays the signals on Probe 1 and Probe 2, where there appears to be a low-frequency oscillation common to both signals.

The magnitudes of the floating potential oscillations were an order of magnitude lower than the magnitude of potential oscillations measured by Goebel using floating emissive probes [1]. By the nature of their operation, floating emissive probes record a steady-state potential that is closer to the actual plasma potential; however, both floating Langmuir and floating emissive probes should capture the amplitude and frequency of oscillations in the plasma potential. One possible explanation for the difference in amplitude of the oscillations is that the assumptions used in the traditional analysis of data collected with floating Langmuir probes is inappropriate. Another possible explanation could be the different geometric design of the model thrusters and the different operating conditions between experiments; however, it is unlikely that these differences alone could explain the large discrepancy between oscillation amplitude.

Task 4: Simultaneously monitor the oscillations in the discharge plasma and the ion energies.

Instead of sweeping through transmission energies with the ion filters as would be done to obtain ion energy distributions, the energy-to-charge filter was set to a constant transmission energy. Fluctuations in the ion count were monitored and recorded along with fluctuations in the floating potential of the plasma within in the detector's field of view. The intensity of the current signal associated with ions of energies near the discharge voltage followed the oscillations in the floating potential of the plasma.

The amplitude of the oscillations in the plasma potential when added to the DC plasma potentials is insufficient to accelerate the ions to the observed energies. The oscillations measured in this experiment were only several volts. This measured amplitude was substantially lower than the amplitudes measured in previous experiments with different diagnostics; Goebel [1] presented data that show peaks up to 45 V in plasma potential oscillations. However, there are no measurements of plasma potential oscillations that could explain the 85-V peak in the ion energy distribution, nor the presence of the high-energy ion tail.

Figure 3 displays the relationship between the floating probe signal and the energy analyzer signal when the probe is located within the analyzer's field of view on the centerline of the thruster. Though there are oscillations at several frequencies that are not common to both signals, there is an apparent trend between the signals. Interestingly, the delay between features in the signals does not agree with the expected delay based on the ion speed. For example, in Figure 3, the velocity of 200-V ions is an order of magnitude lower than the velocity required for an ion to travel between the probe and the energy analyzer in the time delay that is indicated in the figure.

C. SIGNIFICANCE OF RESULTS

Though oscillations in the plasma potential may explain peaks in the ion energy distributions that are nearly double the discharge voltage, these oscillations cannot explain the peaks at higher energies or the tail of the distribution. The observed relationship between oscillations in the floating potential of the plasma and ion energies provides direct evidence that a wave-based acceleration mechanism is responsible for imparting energy to the higher-energy ions. Though several researchers, both at JPL and outside of JPL, have observed ions with anomalously high energies and have hypothesized wave-based acceleration mechanisms, the results of this current study provide the first direct evidence to support such a hypothesis. However, the discrepancy in the expected time delay between the correlated signals must be understood before a conclusion can be reached regarding the acceleration processes.

Erosion of the hollow cathode assembly in ion engines is a major life-limiting mechanism; significant erosion was noted during the life testing of the NSTAR flight spare engine. Understanding the mechanisms that lead to production of the high-energy ions that bombard the surfaces, leading to erosion, can lead to design improvements that reduce erosion and improve thruster life. In addition, existing models of thruster performance can be updated to better predict ion engine life.

D. NEW TECHNOLOGY

No new technology resulted from this DRDF task research.

E. FINANCIAL STATUS

The total funding for this task was $93,458, all of which has been expended.

F. ACKNOWLEDGEMENTS

The authors would like to recognize Al Owens and Ray Swindlehurst for sharing their technical knowledge to assist in implementing the experiments. In addition, the authors would like to thank Dr. Paul Bellan for sharing his expertise on wave phenomena in plasmas.

G. PUBLICATIONS

None

H. REFERENCES

[1]   Dan M. Goebel, Kristina K. Jameson, Ira Katz, and Ioannis G. Mikellides, "Potential Fluctuations and Energetic Ion Production in Hollow Cathode Discharges," *Physics of Plasmas* 14 (October 2007), p. 103508.

I. FIGURES
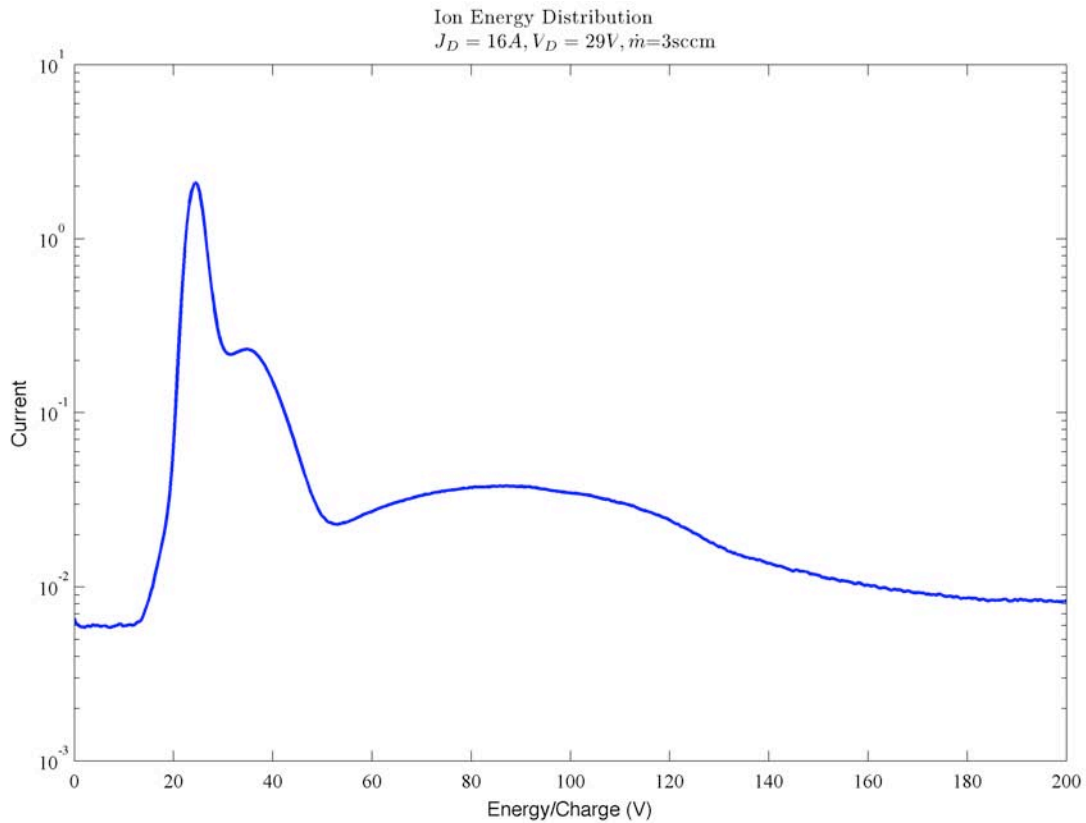
Figure 1. The ion energy distribution obtained when the discharge conditions were set to 16 A and 3 sccm shows three distinct peaks. The first peak at 24 V is associated with the discharge voltage of ~29 V. There are two peaks in the distribution at 35 V and at 85 V, both above the discharge voltage. In addition, the tail in the ion energy distribution shows ions with energies-per-charge in excess of 200 V.
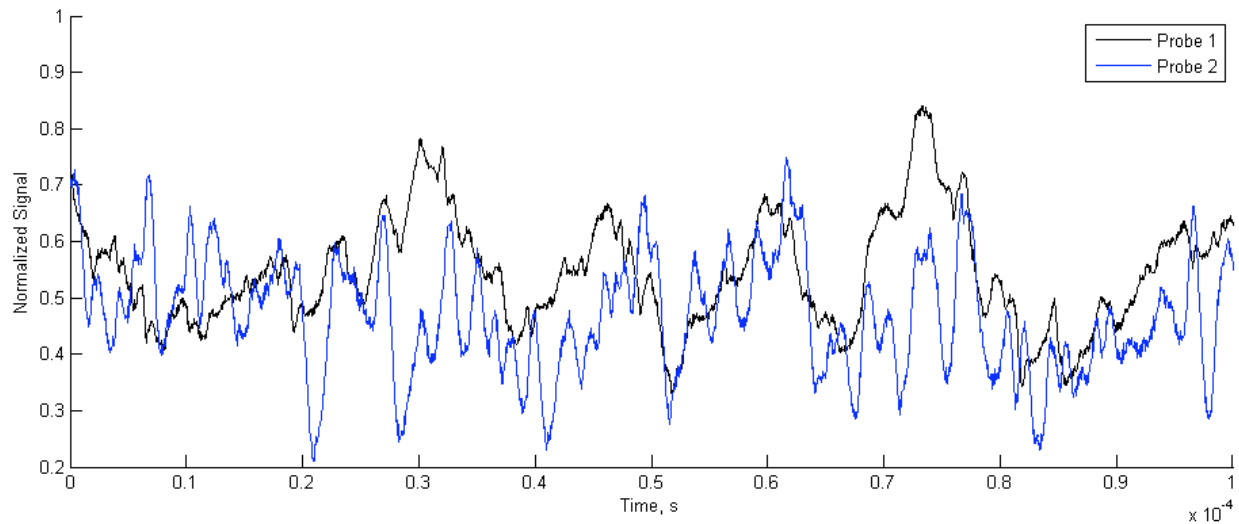
Figure 2. Probe 1 and Probe 2 show similar trends in low-frequency oscillatory behavior. When the signal on probe 1 dips, it is followed by the signal on Probe 2 dipping. The position of Probe 2 while these data were collected was r = 1.5 cm and z = 2.5 cm. The absolute distance of Probe 2 from the thruster centerline is greater than the absolute distance of Probe 1.
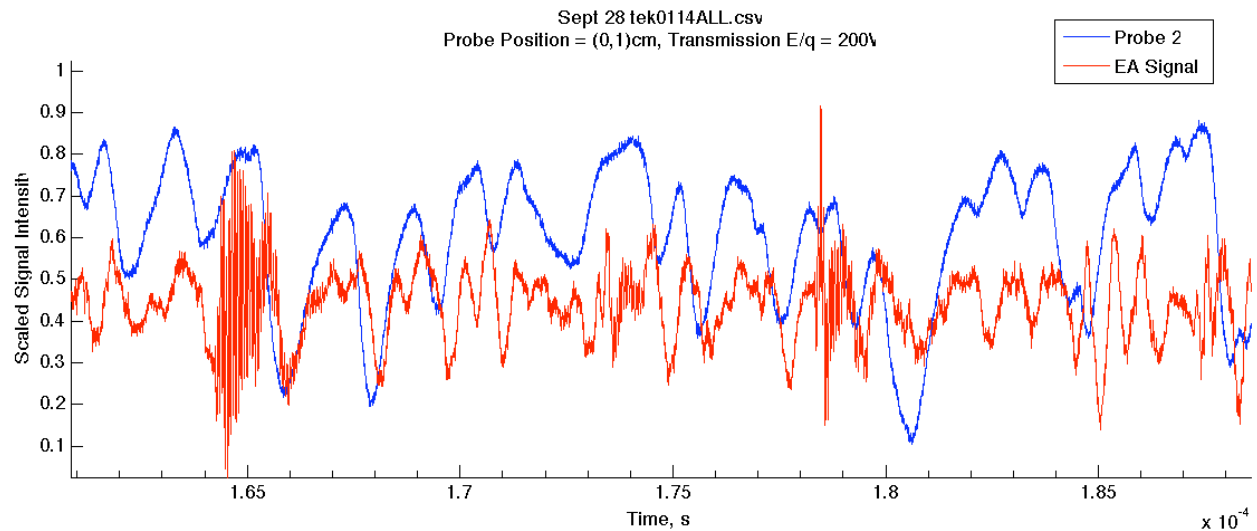
Figure 3. Though the signal on the energy analyzer does not follow the signal on the moveable probe exactly, there are several features that appear in both signals. In particular, the dips in the energy analyzer signal tend to follow dips in the probe signal by ~4e-7 s. For the time series displayed above, the energy analyzer transmission energy-per-charge was 200 V, which is approximately seven times greater than the discharge voltage.

ROBOTICS AND AUTOMATION

DYNAMIC STABILITY FOR EFFICIENT LEGGED LOCOMOTION

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1325

Christopher Assad (PI), Mobility and Robotic Systems Section (347)
Mitra J.Z. Hartmann (Co-PI), Mechanical and Biomedical Engineering, Northwestern University
Joseph H. Solomon, Mechanical Engineering, Northwestern University
Mark A. Locascio, Mechanical Engineering, Northwestern University
Tuomas Haarnoja, JPL Graduate Student Intern, Mobility and Robotic Systems Section (347),
and Aalto University, Helsinki, Finland
Rudranarayan M. Mukherjee, Mobility and Robotic Systems Section (347)

A.   OBJECTIVES

The goal of this research was to develop tools and control methods to help achieve dynamically stable bipedal walking. Legged locomotion has significant potential advantages for NASA missions that involve planetary surface exploration, mobility over rough terrain, or humanoid robot assistants. Such capable legged robots do not yet exist — progress has been impeded by hurdles in power efficiency, as well as in stability, balance, and gait modulation to accommodate for surface terrain and walking transitions. Research on power efficiency has recently achieved a breakthrough with walking bipeds based on "passive dynamics" that attain close to human power efficiency [1,2]. But the stability problem is not yet solved — these bipeds fall after several meters because there is no active control to reject disturbances. Our long-term goal is to develop an active balancing system for passive-dynamic walking robots.

The research had three related objectives:

1)   Develop computer models to simulate the full dynamics of a biped during walking trials. The goal was to generate training databases to learn balance and to test walking control methods.

2)   Develop state estimation and control algorithms for stable and power-efficient walking, first on flat surfaces and then extended to walking over rough terrain. The goal was to perform these steps in simulation before transferring the control methods to a hardware testbed.

3)   Adapt the simulation results, control, and training methods for a power-efficient bipedal robot platform based on passive–dynamic design principles. Our task goal was to demonstrate stable walking over flat ground with a hardware testbed. The longer-term goal is to develop control methods to walk over rough terrain, including gait transitions and turn maneuvers.

B.   APPROACH AND RESULTS

This research consisted of three components: programming of dynamic walking simulations, evaluation of state estimation and walking control algorithms, and application to a hardware platform, the Northwestern University (NU) Hermes robot.

1.   Dynamic Simulations

Our initial approach was to construct computer models to simulate the full dynamics of the biped during walking and to generate databases for learning to control balance. We built several 2D and 3D biped models and collected data during simulated walking trials. Models were based on three methods: the Open Dynamics Engine (ODE), an open-source software package for simulating the dynamics of articulated rigid body systems; a TMT method for constrained dynamics [3] written at NU; and software developed at JPL for highly efficient multibody dynamics simulations with adaptive temporal integration for hybrid systems [4].

The first models were simplified 2D bipeds with 4 links, as illustrated in Figure 1 (left). This model has point feet and a point mass at the hip. The swing leg is allowed to rotate freely at the hip and knee, while the stance leg has its knee locked. These simplifications allowed preliminary development of state estimation and control algorithms. The NU team then constructed a more sophisticated 2D model (Figure 1, middle) to better approximate the dynamics of the NU Hermes biped. This model omits lateral dynamics, but otherwise reasonably reflects Hermes' geometry, mass distribution, and sensing/control capabilities. It has 7 links: the upper-body, two thighs, two shanks, and two feet. Joint angle limits and contact with the terrain are modeled using "soft" constraints, equivalent to stiff springs and dampers. A traditional Coulomb friction model incorporated both static and dynamic friction coefficients.

The Hermes model was extended to 3D, shown in Figure 2 (right), with parameters selected so that the geometrical and inertial properties of the model match those of the Hermes hardware as closely as possible. Joint angle limits are modeled as rigid constraints using ODE's native joint angle stop functionality, and the ground contact again simulates stiff springs and dampers, while incorporating an efficient friction model.

Rough terrain was modeled with a smoothed "random walk" function that determines morphological features by a few intuitive parameters. The slope ranged from –10º to +10º, and the standard deviation of the step-to-step difference in height was ~ 6% of leg length. Sample 2D terrain is shown in Figure 3 (bottom left), and 3D terrain patches are shown in Figure 4 (left).

2.   Algorithms for state estimation and walking control

a.   Cerebellar Dynamic State Estimation

The JPL team adapted a neural network for nonlinear dynamic state estimation (Figure 5, left) [5] inspired by a functional model of the cerebellum, the brain's "engine of agility." This neural net system can enable predictive control for dynamic balancing throughout the biped's walking cycle by (1) learning to predict periodic trajectories in state space from multiple walking trials, (2) estimating when the system begins to diverge from good

trajectories, and then (3) modifying motor actuation to return to stable trajectories before the divergence grows large.

The network was trained to predict impending falls through an iterative learning process. Multiple-state trajectory datasets were collected from the 4-link 2D biped model during successful walking trials, to learn nominal trajectories, and during failed trials, to learn regions of state space that lead to falls (Figure 3). Through a combination of Hebbian association and reinforcement learning, the network was trained to estimate critical dynamic state variables, track trajectories in state space, and predict falling several steps in advance (Figure 5, right). These results were promising, so in ongoing work we are applying this method to the more realistic, higher degree-of-freedom (DOF) models. Our long-term goal is to implement the trained neural net in the hardware testbed to achieve real-time walking balance.

b.   The Braitenberg Biped Controller

For the 7-link Hermes biped models, the NU team developed a "Braitenberg locomotion controller" that controls the joints with a two-layer, feedforward neural network. This neural net has direct linear connections from each of N input nodes, which represent sensor states, to each of M output nodes, which represent desired angles for the associated joints. The desired angles were then fed into standard proportional-derivative controllers to compute the applied torques. This control scheme was applied to control of both walking and standing. For walking, a simple state machine was also used to determine which leg was the stance leg and which the swing leg.

The Braitenberg controller requires over 200 parameters to be set properly for stable locomotion to emerge. To optimize parameters, an evolutionary algorithm (EA) was written to maximize the desired fitness function: the distance the robot could walk before falling down, or before using 1000 J of energy. A total of 14 EA runs were conducted on the 2D model, using a population size of 150, each run lasting 1000 generations. Stable walking controllers evolved for each run, including gait initiation behaviors. Given the 1000 J of energy to consume, the average controller could walk 75.7 m, and only fell on 7% of the terrains. Four EA runs were also conducted on the 3D ODE model of Hermes. Each run evolved controllers that were consistently able to traverse the rough terrain without falling (Figure 4). Although we are in the process of producing more runs and analyzing the data, it is clear that the performance achieved in simulation would be a breakthrough if carried over to the actual robot. The biped simulation shows impressive ability to recover from the step-ups and step-downs in the terrain.

3.   Hardware

The NU limit cycle walking biped "Hermes" (Figure 2) is a copy of the "Flame" robot designed and fabricated in the lab of Prof. Martijn Wisse and colleagues at TU Delft, Netherlands [6]. Hermes is fully 3D in that lateral dynamics are not constrained. It possesses nine internal DOFs: one to control lateral interleg spacing, one at each hip, one at each knee, and two at each ankle. All DOFs are controllable except for the ankle roll. The mechanics are designed to exploit its passive dynamics during gait. Hermes' other components include a PC104 computer stack, lithium–polymer batteries, DC motors powering series elastic actuators, joint

encoders, an inertial measurement unit (IMU), and foot switches to detect ground contact. The encoders and IMU together provide the computer with a complete state description of the robot.

In order to minimize the differences between simulation and hardware, the NU controller code was designed to seamlessly switch between ODE and "real-life physics" modes. That is, the controller calculates the torques to apply with each motor. Then, it can either apply those torques to the joints of the simulated robot in ODE, or send them to the actual physical robot. After a short time interval (1 ms), the joint angles and velocities are collected either from the physical robot's sensors or from the ODE model, and the process repeats. Using ODE as a drop-in replacement for actual physics allowed us to run simulations using the same software as was used on the robot, minimizing the risk of divergence.

Before the robot can walk, it first has to stand. Therefore, a Braitenberg standing controller was evolved using the ODE model, with an appropriate fitness function. During each run, random external pushes were applied every half-second to the torso center of mass. The best-evolved controller was successfully transferred to the Hermes robot, which is now able to reject small disturbances and withstand small pushes to the torso while standing (Figure 6).

C.   SIGNIFICANCE OF RESULTS

Our task goals were to demonstrate stable walking over rough ground in simulation, and over flat ground with the hardware platform. We have made significant progress towards those goals and expect to effectively transfer the stable walking controllers to the Hermes robot in ongoing work. The tools and algorithms developed in this task are the first steps towards our long-term goal to develop an active balancing system for dynamically stable locomotion of a biped robot. They will help overcome critical technical hurdles in predictive control of dynamic walking, and should help enable new, improved forms of robotic legged mobility.

Legged locomotion has large potential advantages over wheeled vehicles for NASA missions to explore planetary surfaces with rough terrain, if the stability and power challenges can be overcome. For example, a legged humanoid could operate effectively in and around a surface base designed for humans; a robotic mountain goat could explore craters with steep walls on Mars. Future construction of facilities on the Moon or Mars may also benefit from coordinated teams of humans and humanoid robotic assistants. Bipedal research is also relevant to improving prosthetics for the disabled, and so may benefit society outside of NASA.

D.   NEW TECHNOLOGY

This work was basic research — no new reportable technology has been developed yet.

E.   FINANCIAL STATUS

The total funding for this task was $200,000, of which $198,617 was expended.

F.   ACKNOWLEDGEMENTS

G.   PUBLICATIONS and PRESENTATIONS

[A]   Joseph H. Solomon and Mitra J.Z. Hartmann, "Braitenberg Control of a Seven-Link Walking Model," *Dynamic Walking 2009,* Vancouver, Canada, June 8–11, 2009.

[B]   Christopher Assad, "Cerebellar Balance for Humanoids," *Dynamic Walking 2009,* Vancouver, Canada, June 8–11, 2009.

[C]   Joseph H. Solomon, Mark Locascio, and Mitra J.Z. Hartmann, "A Remarkably Simple Controller for Bipedal Locomotion over Rugged Terrain," *Symposium on Neuromechanics Research*, Chicago, Illinois, May 17–18, 2010.

H.   REFERENCES

[1   Tad McGeer, "Passive Dynamic Walking (Two-Legged Machines)," *International Journal of Robotics Research* **9** (1990): pp. 62–82.

[2]   Steven Collins, Andrew Ruina, Russ Tedrake, and Martijn Wisse, "Efficient Bipedal Robots Based on Passive–Dynamic Walkers," *Science* **307** (February 18, 2005): pp. 1082–1085.

[3]   Martijn Wisse, "Essentials of Dynamic Walking: Analysis and Design of Two-Legged Robots," Ph.D. Thesis, TU Delft, 2004.

[4]   Rudranarayan M. K. Mukherjee and K. S. Anderson, "Orthogonal Complement Based Divide-and-Conquer Algorithm for Constrained Multibody Systems," *Nonlinear Dynamics* **48** (2007): pp. 199–215.

[5]   Christopher Assad, Salomon Trujillo, Sanjay Dastoor, and Ling Xu, "Cerebellar Dynamic State Estimation for a Biomorphic Robot Arm," IEEE Int. Conf. Systems, Man, and Cybernetics, pp. 877–882, Waikoloa, Hawaii, Oct 10–12, 2005.

[6]   Daan Hobbelen, "Limit Cycle Walking," Ph.D. Thesis, TU Delft, 2008.

I. FIGURES
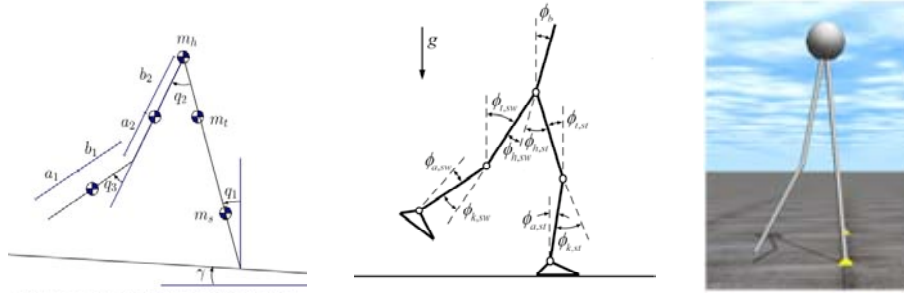
Figure 1. DOF diagrams of 2D biped models with 4 and 7 links, and ODE visualization.
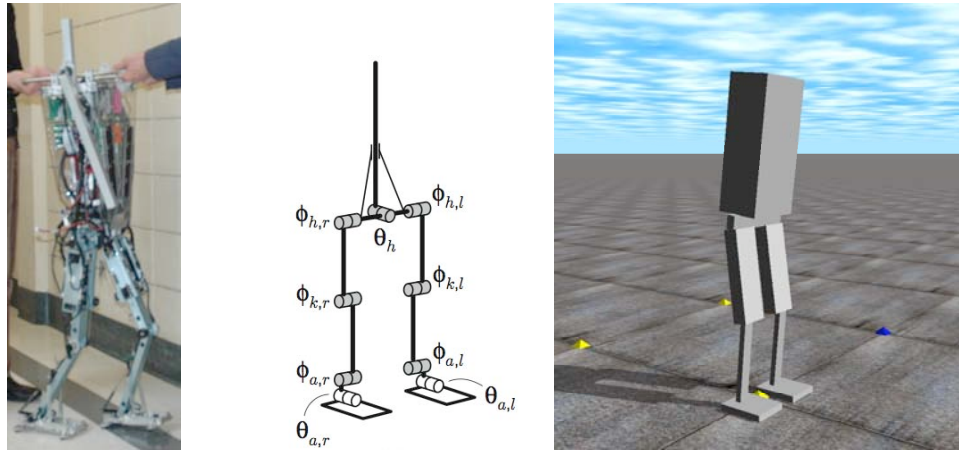
Figure 2. The NU Hermes biped, DOF diagram (from [6]), and 3D model in ODE.
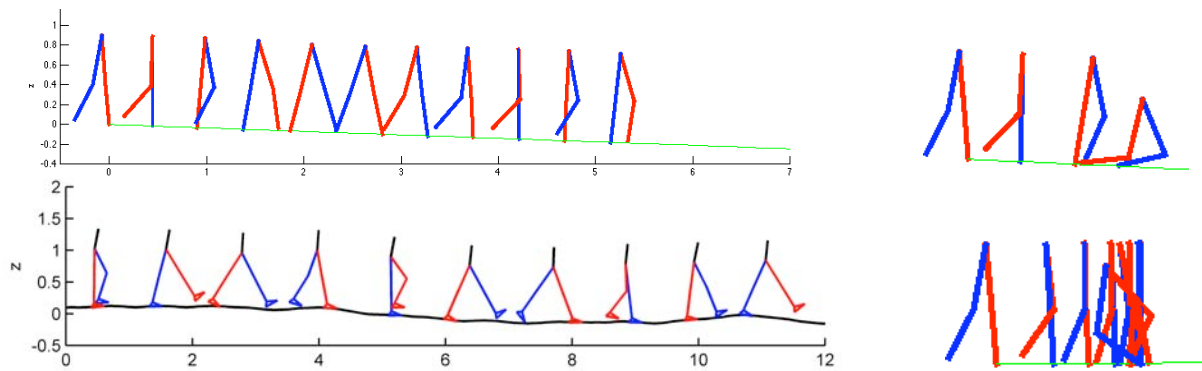
Figure 3. Left: 2D model simulations. Right: Forward falls typically involve scuffing the swing foot or not extending the swing leg fast enough; backward falls occur on slightly negative slopes.
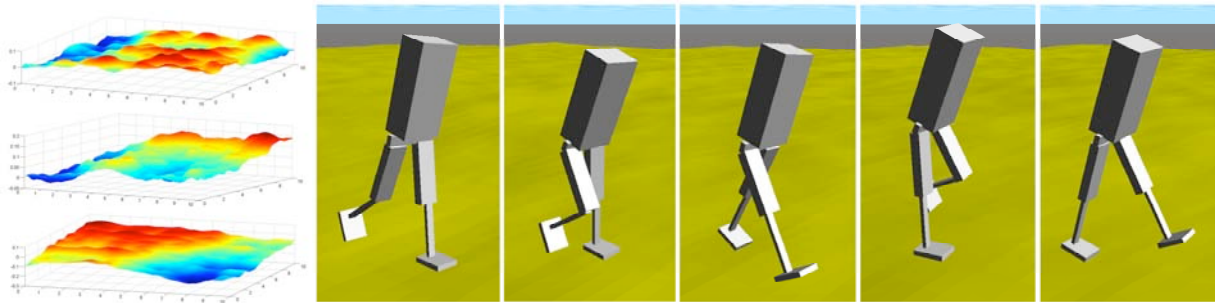
Figure 4. 3-D simulation results. Left: Randomly generated rough terrain patches. Right: ODE model of Hermes successfully walking on rough terrain with an evolved Braitenberg controller.
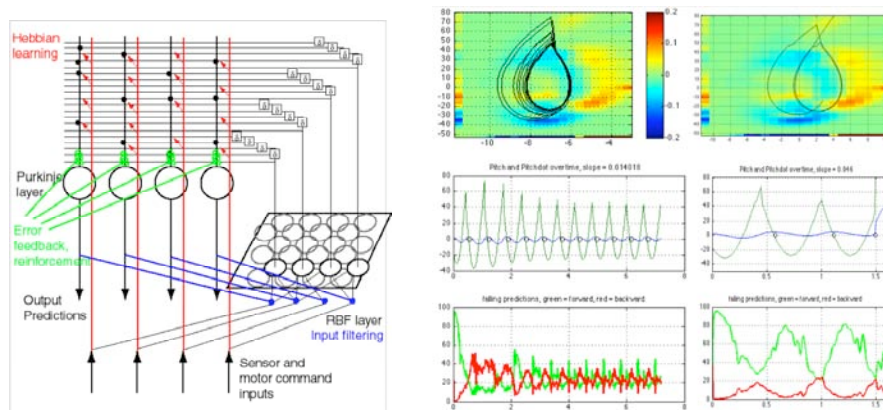
Figure 5. Left: Cerebellar neural network. Right: State space trajectories and falling predictions from walking trials of the 4-link model. Colors at top represent network weights after learning.
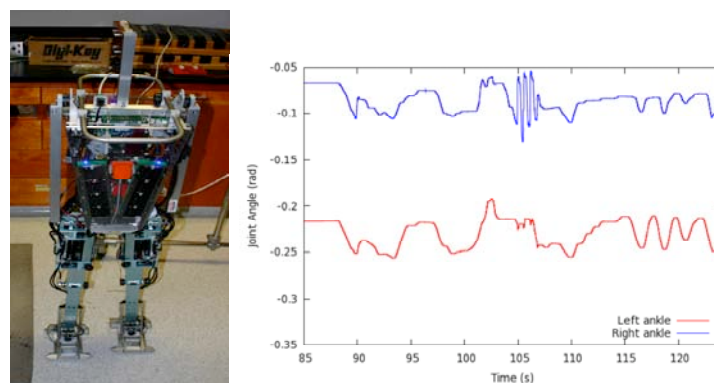
Figure 6. Left: Hermes in standing balance control mode. Right: The joint angles plotted over time show Hermes is able to effectively reject small push disturbances and remain standing.

SCIENCE

CLIMATE MODEL TESTING WITH GPS RADIO OCCULTATION DATA

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1322

Chi O. Ao (PI), Tracking Systems and Applications Section (335)
James G. Anderson (Co-PI), Harvard School of Engineering
and Applied Science, Harvard University
Stephen S. Leroy, Harvard School of Engineering and Applied Science,
Harvard University

A.   OBJECTIVES

GPS radio occultation (RO) measurements have been known to possess characteristics that make them particularly suitable for climate benchmarking [1]. In this study, we explored the application of GPS RO data in climate model testing. In order to facilitate comparisons with climate models, a Bayesian method was first applied to interpolate the quasi-randomly distributed RO measurements onto regular latitude–longitude grids. The 8-year continuous time series of RO measurements from the CHAllenging Minisatellite Payload (CHAMP) and Constellation Observing System for Meteorology, Ionospheric, and Climate (COSMIC) satellite missions were calculated and compared with the European Centre for Medium-range Weather Forecast (ECMWF) reanalysis and the climate models from the Fourth Scientific Assessment of the Intergovernmental Panel for Climate Change (IPCC AR4).

B.   APPROACH AND RESULTS

1.   Data Description

To obtain over 8 years of continuous RO measurements, we used data from both the CHAMP (April 2001–August 2008) and COSMIC (also known as FORMOSAT-3, June 2006–present) missions. The CHAMP mission yielded approximately 150 profiles per day over the globe, with local time sampling that repeats approximately every 4 months. With six satellites capable of performing both rising and setting occultations, COSMIC data provided an order-of-magnitude improvement in throughput as well as excellent diurnal cycle sampling. The period of overlap between the CHAMP and COSMIC data allowed us to assess sampling error from the more limited coverage of the CHAMP data.

The CHAMP and COSMIC data were retrieved at JPL [2]. Each occultation yields a vertical profile of refractivity with high vertical resolution (approximately 200 m up to the lower stratosphere). The refractivity can be expressed as a function of temperature, pressure, and water vapor. By neglecting water vapor, the temperature and pressure can be obtained directly from the hydrostatic equation. The temperature and pressure thus derived are referred to

as the "dry" temperature (Tdry) and "dry" pressure (Pdry). They differ from the physical temperature and pressure only when water vapor is abundant (below about 8 km altitude in the tropics).

2.   Data Gridding with Bayesian Interpolation Method

To perform climate averaging with the GPS RO data and to compare with climate models, it is convenient to first interpolate the quasi-randomly distributed measurements onto a regular latitude–longitude grid. In this study, the gridding was carried out within the framework of Bayesian interpolation [3]. The Bayesian interpolation method uses a set of basis functions to fit irregularly sampled data with unknown noise characteristics and represents a generalized least χ^2 method. What makes this approach Bayesian is its use of a regularizing function that is weighted optimally against data misfit to resolve structures in the data without overfitting. For interpolation on a sphere, we used spherical harmonic functions as the basis functions. The regularizer functions were chosen to be simple power laws that assigned larger penalties to spherical harmonics with larger degrees and non-zonal orders ($m \neq 0$) [3].

Figure 1 illustrates an example of the Bayesian interpolation of GPS RO data. In this example, the temperature data at 10 km height level from January 1–4, 2007, were fitted with spherical harmonic functions up to degree 8. Having obtained the fitting coefficients, the temperature at any latitude and longitude can be easily and quickly computed. Here, the interpolated temperature was calculated on a 10 deg × 10 deg latitude–longitude grid. Figure 1a shows the interpolated COSMIC temperatures in contour, while Figure 1b shows the estimated error from the Bayesian interpolation that characterizes the global misfit. The actual locations of the measurements are shown in dots. Note that the estimated error is larger in the tropics due to the smaller number of measurements present in the tropics. However, regional variability is not accounted for in this formulation, so the error tends to be overestimated in the tropics where the synoptic variability is lower. Figure 2 shows the zonal averages from COSMIC and CHAMP for the same period. The zonal averages can be easily calculated by summing over the spherical harmonic functions with zero order. It can be seen that due to the tenfold increase in the COSMIC measurements, the estimated interpolation error is smaller by a factor of about 3.

To calculate monthly averages, we first divided the data into bins of ~3 days each. We note that bins with too few days would have an insufficient number of measurements that caused overfitting, while bins with too many days would have increased natural variability that caused underfitting. This suggests that there should exist an optimal bin size for a given sampling density. In practice, however, we found that the results were not very sensitive to the choice of the bin size. We applied the Bayesian interpolation for each of the 3-day bins and interpolated the measurements to a gridded dataset. The gridded data from each of the bins within a month were then averaged at each grid point to obtain a monthly average. To better quantify the sampling error, we performed sampling analysis using the ECMWF gridded data. The ECMWF data were first subsampled at the GPS RO times and locations. The subsampled data were then interpolated and compared against the derived monthly average obtained with the full ECMWF data. For COSMIC sampling density, the global temperature RMS difference was found to be on the order of 0.3 K.

3.   Time Series Analysis and Comparison with Climate Models

The time series of the Pdry from GPS RO was analyzed and compared with climate models. The change in the log(Pdry) is directly related to the change in geopotential height [1]. As such, it represents the layer-averaged temperature below a particular altitude. Figure 3 shows the monthly zonally averaged Pdry at 10 km altitude for 50 N and 50 S latitudes. The dry pressure exhibits clear annual cycles that correspond to the warming and cooling of the troposphere. Figure 3 also shows the monthly Pdry anomalies with the annual cycles removed. During the period when data from both CHAMP and COSMIC were available, the results agreed well with each other, despite the fact that COSMIC had about 10 times more profiles. This suggests that CHAMP provided adequate sampling for studying monthly zonal means. For comparison, corresponding results from ERA-interim are also shown in Figure 3. The agreement between GPS RO and ERA-interim is quite good.

We performed linear regression analysis on the log(Pdry) anomalies in the period of 2001–2009 (CHAMP from 2001–2006 and COSMIC from 2007–2009). Figure 4 shows the linear trends obtained from GPS RO data for 50 N and 50 S at 10 km height level and how these trends compare with ERA-interim and an ensemble of IPCC AR4 climate models with SRES A1B greenhouse gas forcing scenario. These results show that (1) there are broad agreements between GPS RO, ERA-interim, and most of the models, and (2) the trends obtained this period are not statistically significant.

Figure 5 shows the global annual average of log(Pdry) from 2002–2009 at 20 km height level. The approximate pressure scale height of 5000 m has been multiplied to convert log(Pdry) to layer thickness (cf. Eq. 6 of [1]). In addition to simple global averaging, we have also computed an "optimized" average using the optimal fingerprinting technique ([1] and references therein), where the climate models had been used to infer the natural variability. In this case, there appears to be very little difference between the simple and optimized averages. To relate the changes in the troposphere thickness with surface temperature, we included the Hadley Center/University of East Anglia Climate Research Unit (HadCRUT3) [4] temperature data record from 2000–2009 in Figure 5. These results showed an interesting divergence between the troposphere thickness and surface temperature in the last few years. Further analysis on the GPS RO interpolation and sampling errors should be carried out in the future to assess the statistical significance of such differences.

C.   SIGNIFICANCE OF RESULTS

This study demonstrated the potential values of GPS RO in climate model testing. We developed and implemented an approach to grid the GPS RO retrieved profiles to facilitate comparisons with models. The tools and analyses developed in this study will be useful not only in climate model testing but also in delineating the sampling accuracy and characteristics of GPS RO missions that are currently in planning, such as CLARREO and other National Research Council Decadal Survey missions, where NASA/JPL is expected to be actively involved.

D.   FINANCIAL STATUS

E.   ACKNOWLEDGEMENTS

F.   PUBLICATIONS

None

G.   REFERENCES

[1]   S. S. Leroy, J. G. Anderson, and J. A. Dykema, "Testing Climate Models Using GPS Radio Occultation: A Sensitivity Analysis," *J. Geophys. Res*. **111**, (2006): D17105. doi:10.1029/2005JD006145

[2]   G. A. Hajj, E. R. Kursinski, L. J. Romans, W. I. Bertiger, and S. S. Leroy, "A Technical Description of Atmospheric Sounding by GPS Occultation," *J. Atmos. Solar-Terr. Physics* **64** (2000): pp. 451–469.

[3]   S. S. Leroy, "Measurement of Geopotential Heights by GPS Radio Occultation," *J. Geophys. Res*. **102** (1997): pp. 6971–6986.

[4]   P. Brohan, J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, "Uncertainty Estimates in Regional and Global Observed Temperature Changes: A New Dataset from 1850," *J. Geophys. Res.* **111** (2006): p. D12106. doi:10.1029/2005JD006548

H.   FIGURES

(a) Interpolated temperature [K]

(b) Estimated interpolation error [K]
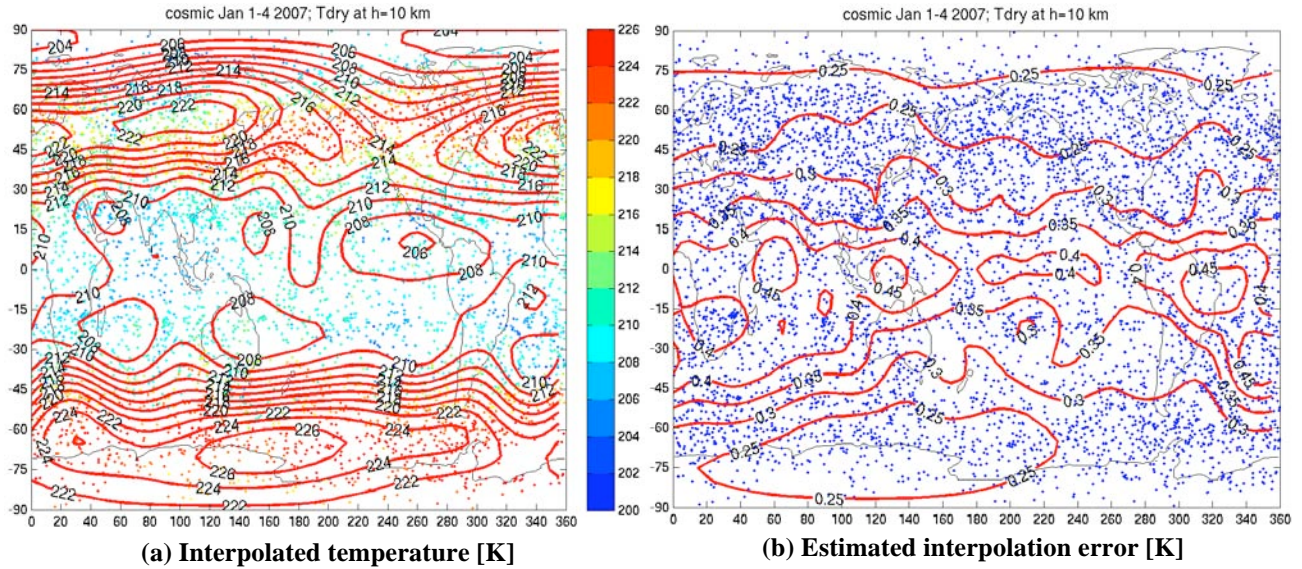
Figure 1. Bayesian interpolation of COSMIC temperature data at 10 km from January 1–4, 2007. (a) Contours from the interpolated temperature. The dots refer to the individual occultation locations and temperature values with color scales shown on the right. (b) Estimated interpolation error. The dots refer to the individual occultation locations.

(a) Zonal mean from COSMIC [K]

(b) Zonal mean from CHAMP [K]
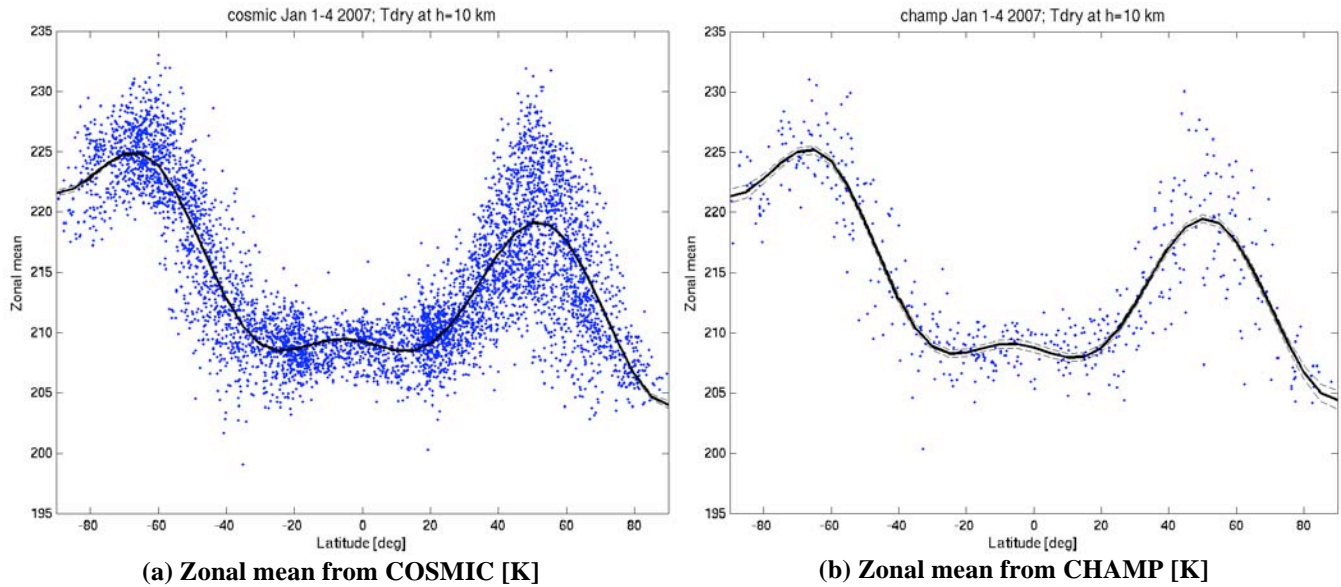
Figure 2. Zonal averages of GPS RO temperature from (a) CHAMP and (b) COSMIC at 10 km from January 1–4, 2007. Solid lines represent the interpolated temperatures and dashed lines represent the estimated interpolation errors. Dots show the actual occultation data.
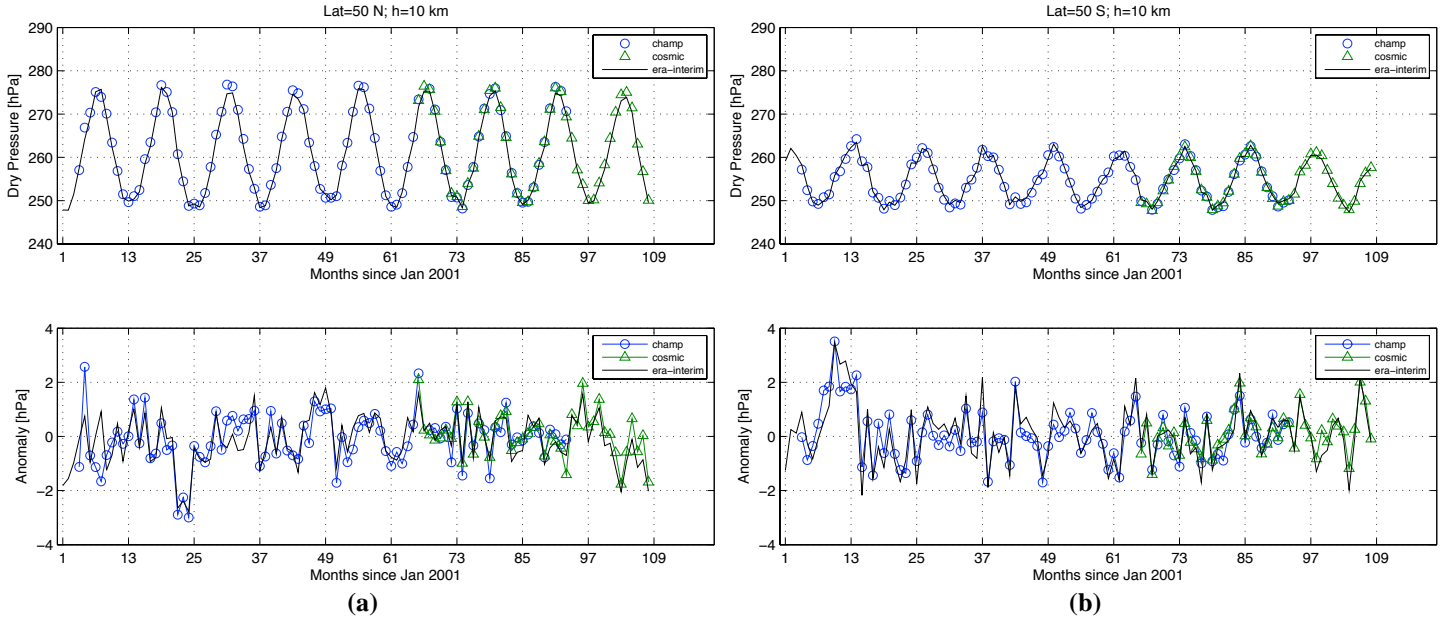
Figure 3. Time series of monthly zonal means of the dry pressure at 10 km obtained from CHAMP (April 2001-August 2008 with July 2007 missing) and COSMIC (June 2006–Dec. 2009) at (a) 50 N latitude and (b) 50 S latitude. For comparison, the corresponding results from ERA-interim are shown. Lower panels show the monthly dry pressure anomalies obtained by removing the annual cycles.
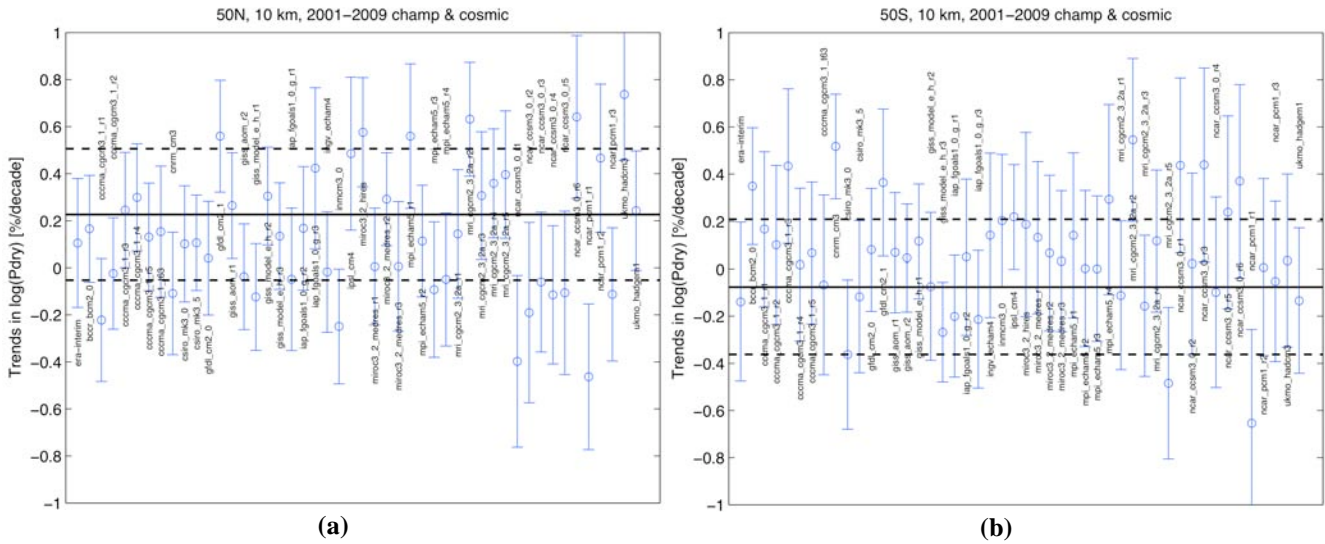
Figure 4. Linear trends in log(Pdry) at 10 km height over 2001–2009 from GPS RO data compared with ERA-interim and IPCC AR4 climate models. (a) 50N; (b) 50 S. Solid black lines denote the trends, while the dashed lines denote the 95% confidence intervals. The corresponding ERA-interim and IPCC model results are shown as circles with error bars denoting the 95% confidence intervals.
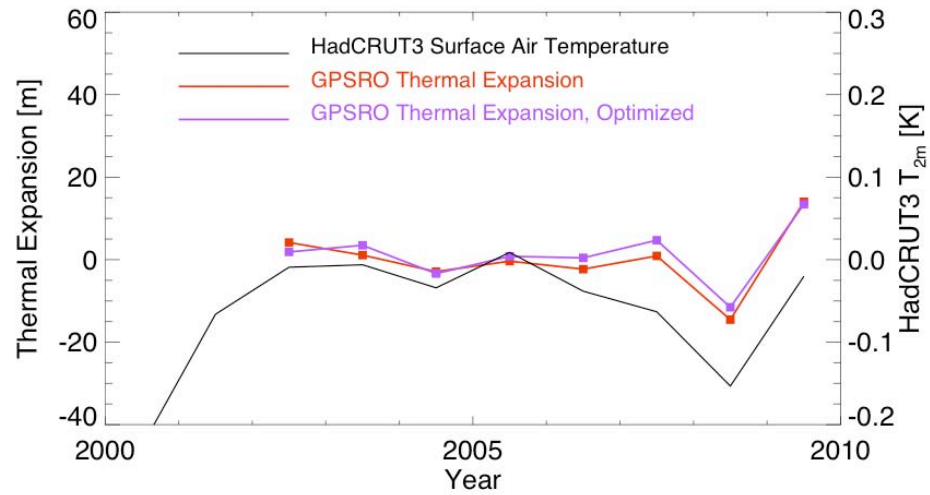
Figure 5. Troposphere thermal expansion inferred from GPS RO log(Pdry) data compared with surface temperature data from HadCRUT3.

ASSIMILATING REMOTE SENSING DATA INTO REGIONAL EARTH SYSTEM MODELS TO ENABLE PREDICTION OF CALIFORNIA ENVIRONMENT IN RESPONSE TO THE PAST AND FUTURE CLIMATE CHANGE

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1331

Yi Chao (PI), Climate, Ocean and Solid Earth Science Section (324)
Zhijin Li, Climate, Ocean and Solid Earth Science Section (324)
Qinbin Li (Co-PI), Department of Atmospheric and Oceanic Sciences and Joint Institute for
Regional Earth System Science and Engineering (JIFRESSE),
University of California at Los Angeles
Dan Chen, Department of Atmospheric and Oceanic Sciences,
University of California at Los Angeles

A.   OBJECTIVES

To address the climate "downscaling" challenge, particularly for California, researchers at UCLA and JPL have established a Joint Institute of Regional Earth System Science and Engineering (JIFRESSE) with a primary goal to develop a state-of-the-art Regional Earth System Model (RESM), which can be used to investigate regional weather and climate processes and help design and test new Earth remote sensing systems. During the past 3 years, JIFRESSE has implemented four interconnected RESM models: the Weather Research and Forecasting (WRF) model for the atmosphere, the Regional Ocean Modeling System (ROMS) for the ocean, the Community Multi-scale Air Quality (CMAQ) model for air quality, and the Simplified Simple Biosphere (SSiB) model for the land–biosphere processes.

Building upon the success of the data assimilation (DA) system for the Regional Ocean Modeling System (ROMS), this DRDF project will implement a DA system for the Weather Research Forecast (WRF) Chemistry (WRF-CHEM) model.

Atmospheric aerosols (also known as particulate matter, PM) are an important air pollutant that has significant human health impacts. The Los Angeles Basin is characterized by poor PM (and ozone) air quality, often the worst in the nation. The goal of our study is to assimilate surface observations into the WRF-Chem model, demonstrating the ability of predicting PM and ozone variability.

There are no comparable efforts within JPL or UCLA. Moreover, to our knowledge there are no comparable efforts outside JPL and UCLA. Further, the remote sensing and data assimilation expertise at JPL complements nicely a greater emphasis on field-based measurements and theoretical modeling at UCLA. To bring data, particularly remote sensing

data, into the developed RESM will provide a rigorous model validation and enable its predictive capabilities.

B.   APPROACH AND RESULTS

As the first step to develop a data assimilation capability for air quality modeling and forecasting, we developed a coupled WRF and chemistry model known as WRF-CHEM over the southern California region (Figure 1). Three one-way nested WRF domains: 36-, 12-, and 4-km and 30 layers in the vertical from the surface to 100 hPa. The 36-km domain covers the western United States with 80×60 grid cells. The 12- and 4-km domains are centered over California and the Los Angeles Basin, respectively.

With the WRF-CHEM, we carried out a detailed evaluation of WRF meteorological simulation for the Los Angeles Basin including surface temperature, wind speed, and direction (Figure 2). It shows good agreement for the day-to-day variation and diurnal cycle of surface (2-meter height) temperature and wind. Understandably, the model has difficulty at simulating near-zero wind speed. Wind speeds were underestimated at the LAX site, suggesting a weaker sea breeze probably due to the active coupling with the coastal ocean.

Using the ARB and IMPROVE surface aerosol observations, we carried out an examination of WRF/CHEM simulation of PM in the LA Basin (Figure 3). It shows the model underestimated by 33% and 57% at the downtown Los Angeles site (N. Main St.) and the Riverside site (Rubidoux), respectively. Model results do not capture the observed diurnal cycles of $PM_{2.5}$ with large underestimation after the morning rush hours. This underestimation could be partly explained by too-high boundary layer heights (dashed blue lines) in the model.

We also improved high-resolution, anthropogenic emissions for WRF/CHEM for the Los Angeles Basin. We carried out a systematic validation of WRF/CHEM ozone chemistry in the Los Angeles Basin using ARB and CalNex data to describe the diurnal cycle, weekday vs. weekend differences, nighttime chemistry, reactive nitrogen (NOy) partitioning.

A conceptual framework to implement the 3-dimensional variational data assimilation (3DVAR) system was developed (Figure 4). A major focus of our future work will be to assimilate surface (and remote sensing) measurements of aerosols into the WRF- CHEM mode for improved $PM_{2.5}$ prediction in the Los Angeles Basin. The data assimilation approach to be used is based on the 3D-VAR assimilation package used in the ROMS coastal ocean model [1,2]. A schematic description of our approach is shown here.

The data assimilation implementation into WRF-CHEM is still in progress and will be carried out by a visiting scientists at JIFRESSE/UCLA funded by a DoE project on a related task. We plan to submit a proposal to the NASA Applied Science Division Human Health and Air Quality programs in response to the ROSES 2011 calls.

C.   SIGNIFICANCE OF RESULTS

This study is the first of its kind in bringing to bear satellite observations to reduce uncertainties in regional air quality prediction. At present, there are no other regional Earth system models that are as comprehensive in terms of the totality of the Earth system processes represented, the targeted resolution (~ O(1 km)), and the level of physical, biological, geographic, and chemical details. There are no other organizational entities, partnered in the manner proposed here between JPL and UCLA, that are in a position to readily apply a comprehensive suite of satellite data to the model development and validation effort, let alone to take advantage of the data assimilation capabilities that enable space-based Earth science mission design and development.

D.   FINANCIAL STATUS

The total funding for this task was $100,000, all of which has been expended.

E.   ACKNOWLEDGEMENTS

F.   PUBLICATIONS

None.

G.   REFERENCES

[1]   Z. Li, Y. Chao, J.C. McWilliams, and K. Ide, "A Three-Dimensional Variational Data Assimilation Scheme for the Regional Ocean Modeling System," *Journal of Atmospheric and Oceanic Technology* **25** (2009): pp. 2074–2090.

[2]   Z. Li, Y. Chao, J. C. McWilliams, and K. Ide, "A Three-Dimensional Variational Data Assimilation Scheme for the Regional Ocean Modeling System: Implementation and Basic Experiments," *Journal of Geophysical Research (Oceans)* **113** (2008): p. C05002.
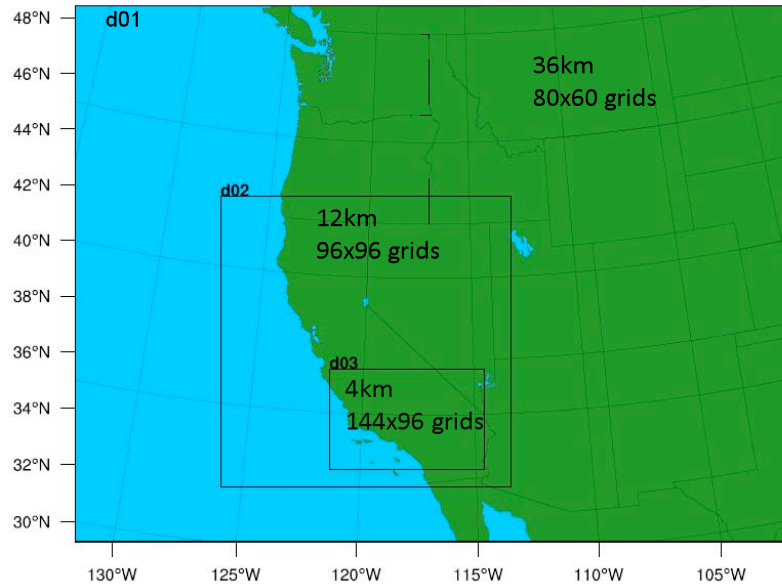
Figure 1. Three one-way nested WRF domains: 36-, 12-, and 4-km and 30 layers in the vertical from the surface to 100 hPa. The 36-km domain covers the western United States with 80×60 grid cells. The 12- and 4-km domains are centered over California and the Los Angeles Basin, respectively.
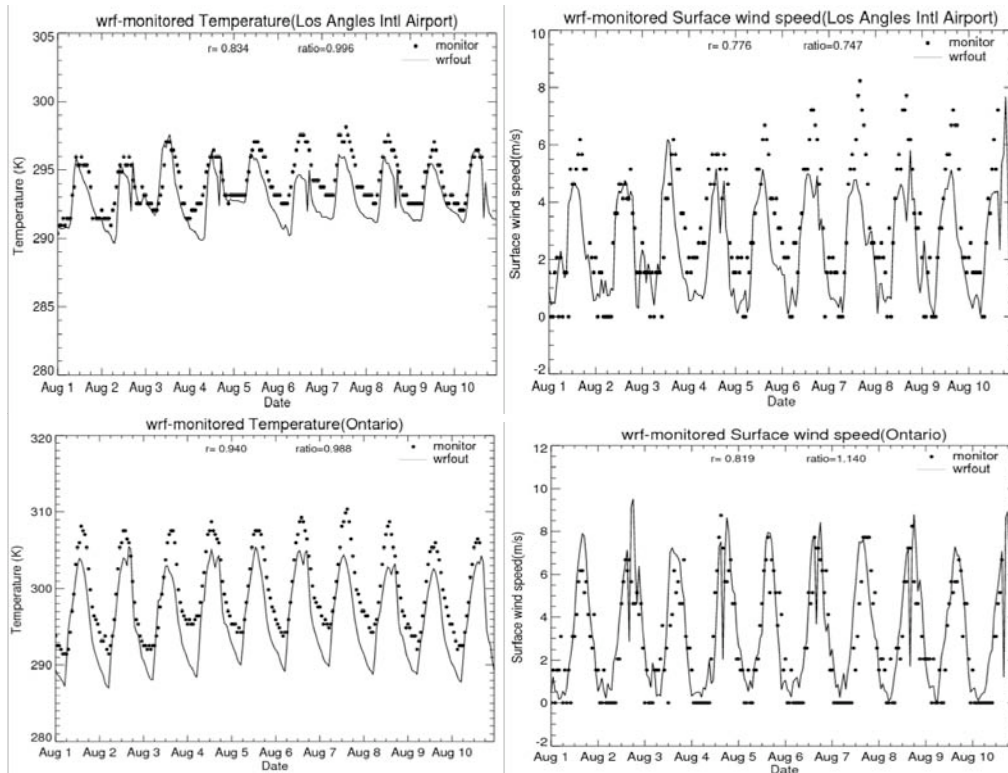
Figure 2. Comparison between observations (dots) and WRF model simulations (lines).
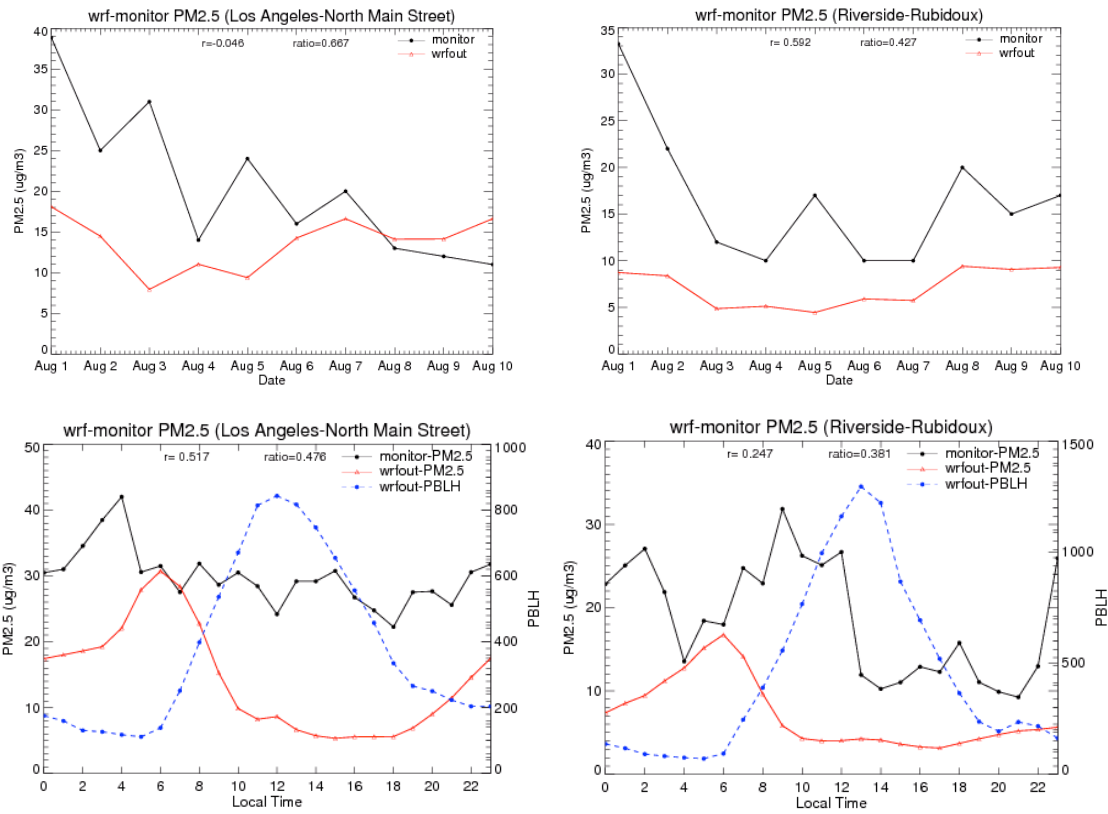
Figure 3. Comparison between observed (black lines) and WRF simulated daily PM$_{2.5}$ (red lines).
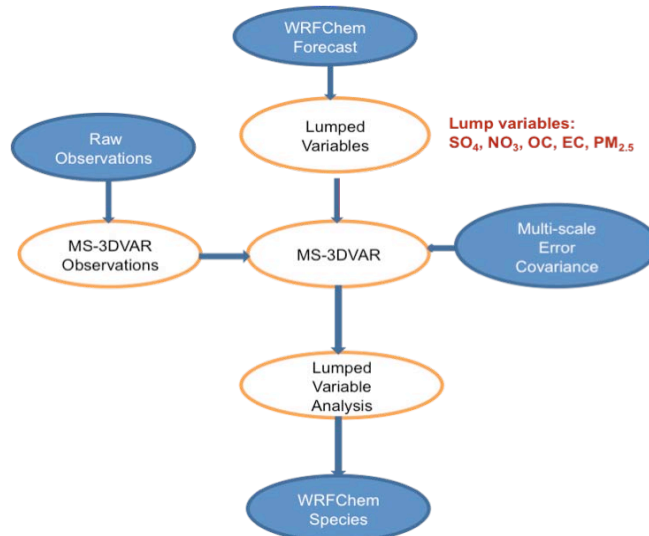
Figure 4. Proposed data assimilation framework for WRF-CHEM.

MONITORING MICRODISCHARGES FROM MARTIAN DUST STORMS DURING DSN DOWNLINKS

Thomas B. H. Kuiper (PI), Astrophysics and Space Sciences Section (326)
Walid A. Majid, Deep Space Tracking Systems Section (335)
Nilton O. Renno (Co-PI), Department of Professor of Atmospheric,
Oceanic and Space Sciences, University of Michigan
David Catling, Department of Atmospheric Sciences, University of Washington
Imke de Pater, Department of Astronomy, University of California at Berkeley

A.   OBJECTIVES

The objective was to demonstrate that the VLBI Science Receivers (VSRs) installed in each Deep Space Network tracking complex could be used to monitor Mars for electrical activity manifested by nonthermal electromagnetic radiation.

The specific goals were to a) develop real-time software for recording nonthermal X- and Ka-band radio emissions from Mars; b) verify the emission in real time with the Allen Telescope Array (ATA); c) use the much wider frequency coverage of the ATA to obtain information about the particle size distribution; and d) request dynamic scheduling of the VLA to localize the emission on Mars; and e) define a set of diagnostic data to be recorded whenever Mars is in the field of view of a DSN antenna.

B.   APPROACH AND RESULTS

1.   Observations

We made observations with the Goldstone DSS-13 antenna and the ATA independently because there was little probability of an electrical event during this period and because the ATA is not yet ready for short-notice dynamic rescheduling.

a.   We used the Goldstone R&D station DSS-13 because of its flexibility in responding to user requirements. The S-/X-band and X-/Ka-band receivers and the VSR were identical to those used at any DSN beam-waveguide 34-m antenna. (Enhanced kurtosis was recorded at S-band with DSS-13 during a convective dust storm on Mars on 2006 June 8 [1].) We obtained 43 nights of data, of typically 12 hours duration, between February 4 and April 23, 2010. We recorded four pairs of adjacent ½-MHz sub-bands, two pairs for each receiver. We computed the mean, variance, skewness, and kurtosis of the signal in each subband in near real time.

b.   The Allen Telescope Array was used to observe Mars on 5 days between March 8, 2010, and April 16, 2010, for approximately 6 hours per day. Observations were conducted using one or more ATA beamformers, with between 20–25 antenna-polarizations phased together in a tied array beam at any one time. When available, two ATA beamformers were used to observe two frequency bands, centered at 3.2 GHz and 8.0 GHz.

A special wideband signal processor was developed for the ATA observations using the tools and infrastructure of the Center for Astronomy Signal Processing and Electronics Research (CASPER) group [2]. This instrument computes the signal moments from a digital stream of complex-baseband voltage data delivered from the ATA beamformer. Each ATA beamformer delivers a 104.8-MHz bandwidth tunable between 500 MHz and 11 GHz. The wideband spectral processor developed for this work calculates both (full-Stokes) power and kurtosis of the input signal as a function of frequency utilizing a 1024-pt polyphase filter bank, in the manner described in [3]. The accumulation time of the instrument is a tunable parameter; for the observations described herein the accumulation time was fixed at 1.25 ms. Accumulated data were delivered via Ethernet to one of several available host PCs.

2.   Results

We achieved goals (a), (c), and (e) described in the Objectives section. Goal (b) was partially achieved.

a.   For the DSN/VSR observations, we developed an extensive set of software for scheduling, launching, and conducting the data acquisition, and reducing the data obtained. The core software is a program that reads the data being written to disk by the VSR and computes the signal statistics. To start the observations, the investigator interacts with a Python script for about a minute and can then log off. No further action is required by the investigator until the end of the observations. The script generates the macro files used to operate the DSS-13 antenna (something that would not be required during a telemetry downlink, at least not by the investigator) and the VSR. Also, command files are generated to initiate the programs that run the antenna, the VSR, and the signal statistics programs at designated times.

A program was written to monitor the output from the signal statistics programs for evidence of changes in the kurtosis, and in particular changes in kurtosis caused by Schumann resonances [1]. It will send a detection report to a remote host listening on a designated port. This was not tested because an analysis by one of the Co-Is (D. Catling) showed that suitable dust storms were not likely during the time we made observations. (A separate report is available.)

After the observations are completed, an investigator runs another Python script that downloads all the relevant log and data files, as well as the scripts and macros for archival purposes, to the data reduction computer. The script also plots the system temperature (Figure 1) and the signal variance and kurtosis (Figures 2 and 3).

In the absence of convective dust storms on Mars, we demonstrated the signal moments calculation by selecting X-band frequencies with spacecraft signals. The Ka-band signals (Figure 3) serve as a reference. Variance is essentially power, except for a scale factor, and this is verified by comparing it to the system temperature (Figure 1). The excess kurtosis (the computed kurtosis – 3) is close to zero, the kurtosis expected from a thermal signal. The slight offset of –0.015, which occurred during every observation, can be attributed to systematic effects in the receivers. (It was also seen at X-band when no spacecraft signals were present.) The spacecraft signals affected the kurtosis in the expected way. Mars Odyssey had a spread spectrum oscillating between 8406.36 and 8406.43 MHz because of Doppler shifts. Indeed, the upper sideband of the 8406-MHz channel shows weak negative kurtosis, as expected for spread spectrum emission. The emission at 8439 MHz consists of strong tones close to the band center and both sidebands show strong negative kurtosis.

b. Initial deployment and engineering tests of the ATA observing pipeline were conducted on February 10, 2010. A variety of diagnostic sources were observed to verify instrument functionality, including Jupiter, the Crab pulsar and several artificial satellites. The kurtosis estimator calculated exactly what is expected for Gaussian signals and has a 3σ variance [3]. As expected, a few frequency channels had a Gaussian-deviant kurtosis value (indicating interference), but the band is largely within the 3σ bounds. The scaling factors and bit-selection employed in the wideband spectrometer introduce scaling differences in the computed power and power^2 (used to derive the kurtosis), and thus recovering the expected Gaussian value confirms that scaling factors and bit-selection are being accounted for correctly.

A kurtosis waterfall plot from an observation of a RADARSAT satellite is shown in Figure 4. Clearly, the satellite transponder should be strongly non-Gaussian, as is exhibited. Analysis of our Mars observations from February–April 2010 is currently underway.

C.   SIGNIFICANCE OF RESULTS

We demonstrated that DSN equipment could be used with little effort to monitor DSN downlink bands for nonthermal emission caused by electrical activity on Mars. In fact, this is being done at the DSN Madrid station by a graduate student (Santa Martinez) to fulfill her academic requirements. Scheduling the observations could, in principle, be fully automated from the DSN 7-day schedule. However, timely review of the data and clearing of the VSR data disk requires prompt investigator attention because a VSR can only store about 15 hours of data.

The kurtosis processor developed by the UC Berkeley team provides a low-cost and much more capable alternative to the VSR for monitoring nonthermal electrical activity on Mars.

The experience in hardware, software, and operations acquired during this project played a key role in the development of an instrument proposal for Mars Trace Gas Orbiter 2016.

A follow-on proposal to the NASA Mars Fundamental Research Program is anticipated.

D.   NEW TECHNOLOGY

The wideband signal processor developed for the ATA observations [3] provides a unique capability for that telescope, and the complete design of the instrument is open-source and available to the CASPER community [2] for use at other observatories. The design can be easily adapted for use in the DSN and, with firmware changes, can perform many other useful scientific functions, adding value to DSN tracking operations.

E.   FINANCIAL STATUS

The total funding for this task was $200,000, all of which has been expended.

F.   ACKNOWLEDGEMENTS

G.   PUBLICATIONS

None yet. Further data analysis intended for publication is underway.

H.   REFERENCES

[1]   Christopher Ruf, Nilton O. Renno, Jasper F. Kok, Etienne Bandelier, Michael J. Sander, Steven Gross, Lyle Skjerve, and Bruce Cantor, "Emission of Non-Thermal Microwave Radiation by a Martian Dust Storm," *Geophysical Research Letters* **36** (2009): p. L13202.

[2]   A. Parsons, D. Backer, D. Werthimer, and M. Wright, "A New Approach to Radio Astronomy Signal Processing: Packet Switched, FPGA-based, Upgradeable, Modular Hardware and Reusable, Platform-Independent Signal Processing Libraries," *Proceedings of the XXXth General Assembly of the International Union of Radio Science*, Boulder, Colorado, 2006.

[3]   G. M. Nita, D. E. Gary, Z. Liu, G. J. Hurford, and S. M. White, "Radio Frequency Interference Excision Using Spectral-Domain Statistics," *Publications of the Astronomical Society of the Pacific* **857** (2007): pp. 805–827.
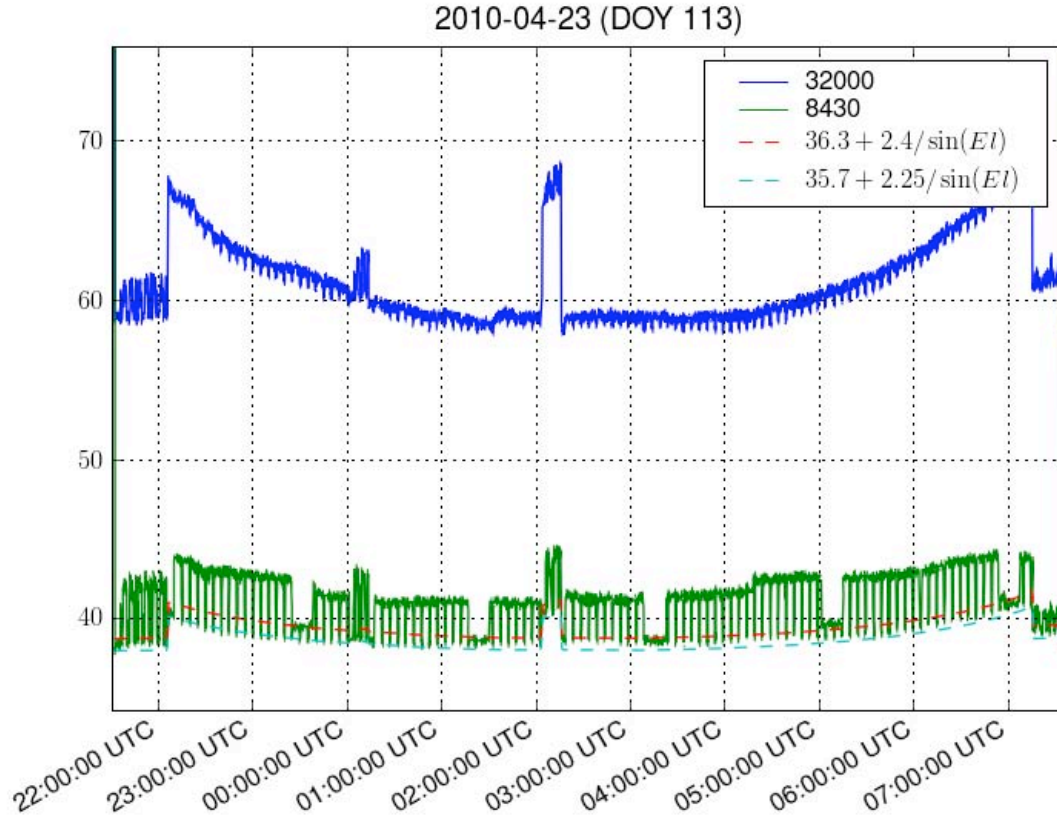
Figure 1. System temperature over 100-MHz bandwidth. Every 5 minutes, the antenna was moved off Mars for 45 seconds. The antenna pointing was checked at the beginning and end and at 0000 UT and 0300 UT.
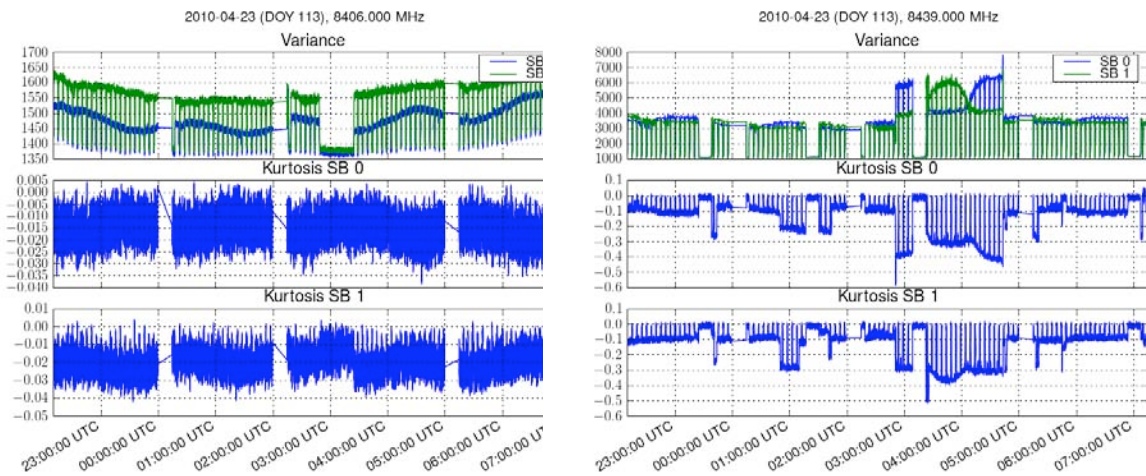
Figure 2. Moments of the signals in the X-band channels. 8406.85 MHz is used by Mars Odyssey. 8439.44 MHz is used by MER-A and MRO.
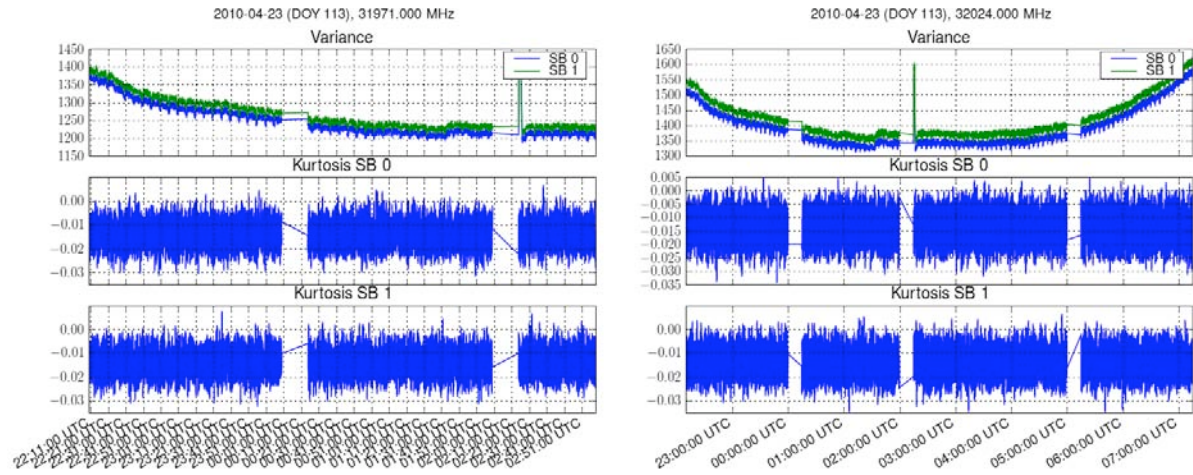
Figure 3. Moments of the signals in the Ka-band channels. The VSR stopped recording the 31971-MHz channel because of a disk recording failure.
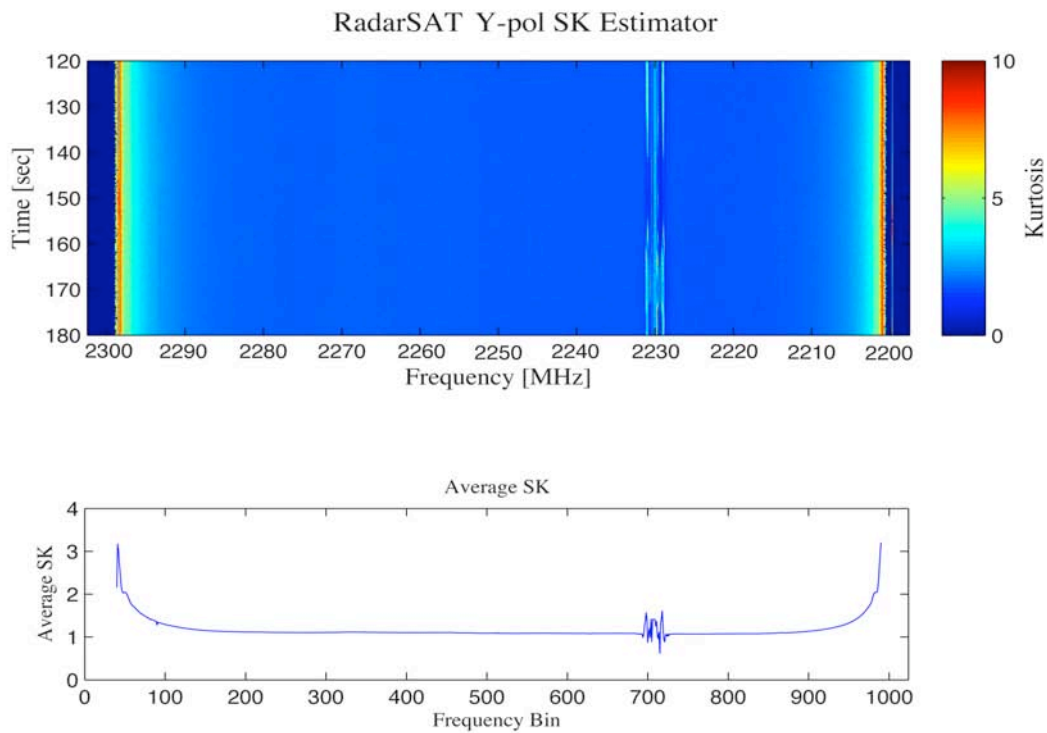
Figure 4. 60 seconds of kurtosis measurements of the 2.23-GHz transmitter of a RADARSAT satellite.

INTERIOR STRUCTURE OF TITAN AND TECTONIC PROCESSES ON TITAN AND EUROPA

Director's Research and Development Fund (DRDF)
Final Report

JPL Task # 1340

Robert T. Pappalardo (PI), Planetary Sciences Section (322)
Giuseppe Mitri, Geological and Planetary Sciences Department,
California Institute of Technology
David J. Stevenson, Geological and Planetary Sciences Department,
California Institute of Technology

A.   OBJECTIVES

Titan is a unique icy satellite not only because it is the only icy satellite in the outer solar system with a substantial atmosphere, but also because Titan presents unique geological characteristics for an icy satellite. Prior to the measurements of the gravity field data of Titan by [1], several studies on the interior structure and evolution of Titan assumed full differentiation of the interior with a deep interior mainly composed of silicate dehydrate [2–5]. Tobie et al. [4] showed that if Titan is fully differentiated with a deep interior mainly composed of dehydrate silicates, then the normalized axial moment of inertia C/MR^2 is <0.33.

Doppler data tracking from four flybys of the Cassini spacecraft was used by [1] to determine the second-degree gravitational constants: $J_2 = (31.808 \pm 0.404) \cdot 10^{-6}$ and $C_{22} = (9.983 \pm 0.039) \cdot 10^{-6}$. Using hydrostatic equilibrium theory with Radau-Darwin approximation, Titan's normalized moment of inertia is found to be $C/MR^2 = 0.3414 \pm 0.0005$ [1], where C is the axial moment of inertia, and M and R are Titan's mass and radius, respectively. Such a relatively high moment of inertia suggests that Titan's deep interior might be formed of relatively low-density material. The fluid Love number of Titan is found to be $k_f = 1.0097 \pm 0.0039$ [1].

Using gravity data as constraints, we model the present interior structure and accretion of Titan, to understand the most plausible interior state of Titan.

B.   APPROACH AND RESULTS

1.   Two- and Three-Layer Models

We consider two- and three-layer models of Titan's interior. The gravitational constant G times mass M_T of Titan is $GM_T = 8.97819 \cdot 10^{12}$ m^3 s^{-2} [6], the mean satellite radius is $R_T = 2,575$ km, and the mean density is 1,881 kg m^{-3}. We use the value of the axial moment of inertia derived from Cassini data as inferred by the gravity coefficients and hydrostatic equilibrium theory [1] as our primary constraint.

The axial moment of inertia gives us an indication of the density distribution of the interior. Figure 1 shows that for $C/MR^2 = 0.34$, a two-layer model predicts that the deep interior has a radius of ~2,100 km and a relatively low density of $\rho_{int} = 2,560$ kg m^{-3}, corresponding to a total mass of the deep interior of M_c ~9.9·10^{22} kg. A three-layer model predicts an even lower density for the intermediate layer; this might be expected if the deep interior hides at its center a silicate or metallic core with a radius of a few hundred kilometers. Iess et al. [1] found that the ratio J_2/C_{22} ~3.186±0.042 is consistent with the theoretical value expected for a body near the hydrostatic equilibrium. However, [1] pointed out that nonhydrostatic contributions to degree-2 gravity coefficients might produce an overestimate of the moment of inertia, and even if the more probable value of C/MR^2 is 0.34, a value of 0.33 must be considered as a lower bound value. Considering this lower bound value, a two-layer model predicts a density of $\rho_{int} = 2,700$ kg m^{-3} and a radius of ~1,990 km for the deep interior (Figure 1).

Two different hypotheses can be advanced to explain such low density: (1) the deep interior is composed of a mixture of ice and rock indicating a partial differentiation, or else (2) it is composed of hydrate silicate, indicating full differentiation as proposed by [7] and pursued by [8].

2.   Partial Differentiation

We examine the accretion and interior structure of a partially differentiated interior. First, we model accretional heating of a growing Titan to consider whether the formation conditions can permit a partially differentiated interior. The gravitational binding energy, $E = (3/5)(GM^2/R_T)$ ~2.8·10^{29} J, gives a first-order estimate of the total amount of energy involved during the accretion of Titan. If all the binding energy is transformed into thermal energy, then the heat produced during the accretion is larger than the required heat to melt Titan's interior ~4·10^{28} J, assuming a latent heat of melting equal to that of the water ice (3·10^5 J kg^{-1}). However, just a fraction of the accretional energy is transformed into retained heat; part of the accretional energy is released from Titan's surface as radiant energy. The release of accretional energy by radiation is facilitated if Titan's accretion occurred on a long timescale.

Radiogenic decay gives a significant contribution to internal heating during the accretion. We determine the radiogenic heating within Titan given by the isotopes ^{238}U, ^{235}U, ^{232}Th, ^{40}K, ^{26}Al and ^{60}Fe. The radiogenic heating for ^{238}U, ^{235}U, ^{232}Th and ^{40}K is determined considering the present abundance within CI chondrite [9]. The start time of our model is the age of calcium-aluminum–rich inclusions in the CV chondrite Efremovka (4.57 Ga) [10].

Figure 2 shows the modeled interior temperature of Titan at the end of accretion. Panel A is for $h = 0$ and $\tau_{accr} = 0.5$ Myr; Panel B is for $h = 0$ and $\tau_{accr} = 1$ Myr; Panel C is for $h = 0.1$ and $\tau_{accr} = 0.5$ Myr; and Panel D is for $h = 0.1$ and $\tau_{accr} = 1$ Myr. All the models are for $t_o = 2.6$ Myr and $t_o = 3.6$ Myr. Figure 2 shows that increasing the start time of accretion (t_0) decreases the total power of heat produced by radiogenic decay, thereby reducing the heating of the interior of Titan. Figure 2 shows that increasing the total accretion time (τ_{accr}) increases the energy

radiated from the surface, which reduces the net accretional heating of Titan. In the models shown in Figure 2, the central temperature is lower than the superficial layers of the deep interior, preventing the onset of convection. However, a comparison of the models of Panels A and B for $h = 0$ and $t_o = 2.6$ Myr shows that for $\tau_{accr} > 1$ Myr, the central temperature might be higher than the upper layers, permitting the onset of convection within the deep interior. Summarizing, we found that Titan can avoid melting of the interior if it was formed relatively late $t_0 \geq 2.6$ Myr, thereby reducing the presence of short-lived radioisotopes, and if the accretion occurred on timescale $\tau_{accr} \geq 0.1$ Myr.

C.   SIGNIFICANCE OF RESULTS

Our work supports the hypothesis that Titan is partially differentiated into a deep interior composed of a mixture of ice and rock. Previous thermal models [2–4]; [11] and [12] suggest that the deep interior is surrounded by a high-pressure (HP) ice layer (with phases III, V, and VI), an ammonia–water subsurface ocean, and an outer ice I shell. Because the deep interior radius does not affect the functional dependence of the depth from the surface of the high ice pressure layer, thermal models developed to describe the outer layers can be altered to a partially differentiated Titan. Figure 3 shows the interior structure of Titan as inferred by thermal models and its moment of inertia.

Titan's normalized moment of inertia ($C/MR^2 \sim 0.34$) indicates that Titan is likely partially differentiated with a deep interior composed of a mixture of ice and rock. We have shown that Titan must be formed relatively late, ≥ 2.6 Myr after the calcium-aluminum–rich inclusions (CAIs), and the accretion must have occurred on a timescale ≥ 0.1 Myr. CI chondrites mineralogical composition for the deep interior of Titan can only be considered if there are present high nonhydrostatic contributions to degree-2 gravity coefficients that would allow $C/MR^2 \sim 0.33$.

Additional gravity data could be used to refine these results; specifically, with additional gravity-dedicated flybys of Titan with the Cassini spacecraft. Given the high demands on Cassini's limited resources, such may require use of the spacecraft's low-gain antenna. Future missions that orbit Titan would be able to refine these results to much greater precision.

D.   NEW TECHNOLOGY

No new technology was expected or emerged from this study.

E.   FINANCIAL STATUS

The total funding for this task was $162,300, all of which has been expended.

F.   ACKNOWLEDGEMENTS

G. PUBLICATIONS

[A] Giuseppe Mitri, "Interior Structure and Accretion of Titan," submitted to *Icarus*.

H. REFERENCES

[1] L. Iess, N. J. Rappaport, R. A. Jacobson, P. Racioppa, D. J. Stevenson, P. Tortora, J. W. Armstrong, and S. W. Asmar, "Gravity Field, Shape, and Moment of Inertia of Titan," *Science* **327** (2010): p. 1367.

[2] O. Grasset, C. Sotin, and F. Deschamps, "On the Internal Structure and Dynamics of Titan," *Planet. Space Sci.* **48** (2000): pp. 617–636.

[3] F. Sohl, H. Hussmann, B. Schwentker, T. Spohn, and R. D. Lorenz, "Interior Structure Models and Tidal Love Numbers of Titan," *Journal of Geophysical Research (Planets)* **108** (2003): p. 5130.

[4] G. Tobie, O. Grasset, J. I. Lunine, A. Mocquet, and C. Sotin, "Titan's Internal Structure Inferred from a Coupled Thermal-Orbital Model," *Icarus* **175** (2005): pp. 496–502.

[5] G. Tobie, J. I. Lunine, and C. Sotin, "Episodic Outgassing as the Origin of Atmospheric Methane on Titan," *Nature* **440** (2006): pp. 61–64.

[6] R. A. Jacobson, P. G. Antreasian, J. J. Bordi, K. E. Criddle, R. Ionasescu, J. B. Jones, R. A. Mackenzie, F. J Pelletier, W. M. Owen, D. C. Roth, J. R. Stauch, and the Cassini Navigation Team, "The Gravity Field of the Saturnian System and the Orbits of the Major Saturnian Satellites," *Bull. Am. Astron. Soc.* **37** (2005): p. 729.

[7] A.D. Fortes, P.M. Grindrod, S. K. Trickett, and L. Vocadlo, "Ammonium Sulfate on Titan: Possible Origin and Role in Cryovolcanism," *Icarus* **188** (2007): pp. 139–153.

[8] J. C. Castillo-Rogez and J. I. Lunine, "Evolution of Titan's Rocky Core Constrained by Cassini Observations," *Geophys. Res. Lett.,* in press. doi:10.1029/2010GL044398

[9] K. Lodders, "Solar System Abundances and Condensation Temperatures of the Eements," *Ap. J.* **591** (2003): pp. 1220–1247.

[10] Y. Amelin, A. N. Krot, E. D. Hutcheon, and A. A. Ulyanov, "Lead Isotopic Ages of Chondrules and Calcium-Aluminum–Rich Inclusions," *Science* **297** (2002): pp. 1678–1683.

[11] G. Mitri and A. P. Showman, "Thermal Convection in ice-I Shells of Titan and Enceladus," *Icarus* **193** (2008): pp. 387–396.

[12] G. Mitri, M. T. Bland, A. P. Showman, J. Radebaugh, B. Stiles, R. M. C. Lopes, J. I. Lunine, and R. T. Pappalardo, "Mountains on Titan: Modeling and Observations," *J. Geophys. Res.* **115** (2010): E10002. doi:10.1029/2010JE003592

I. FIGURES
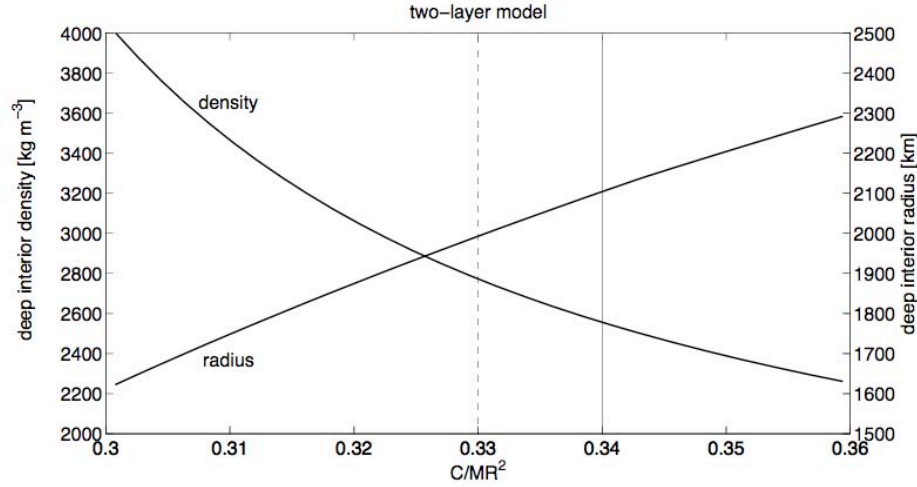
Figure 1. Two-layer model. Deep interior density (left y-axis) and deep interior radius (right y-axis) versus the normalized moment of inertia C/MR^2. Dashed and solid line show plausible range of C/MR^2.
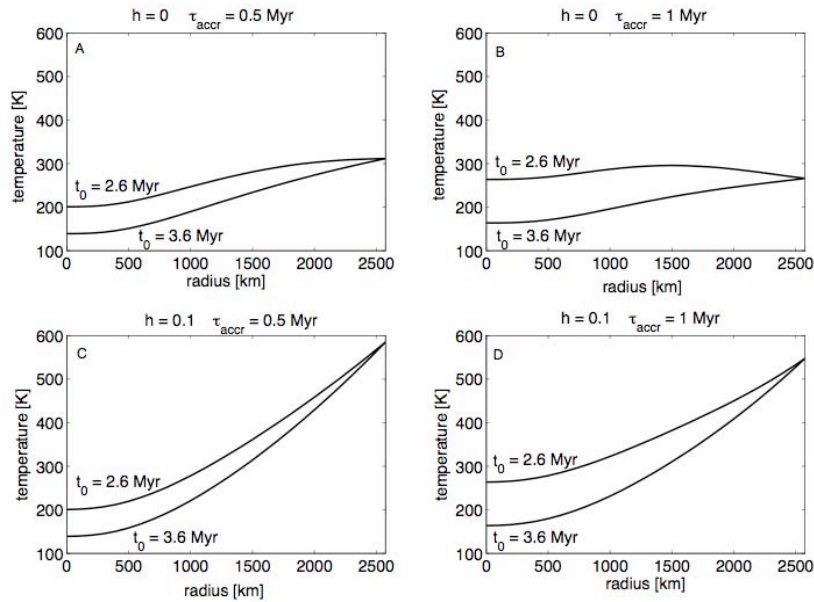
Figure 2. Temperature profile of Titan's interior after accretion for different values of the start time of accretion t_0 and the accretion time τ_{accr}. The time $t = 0$ is the time of the onset of the calcium-aluminum-rich inclusions (CAIs). The model is for $h = 0$ (top panels) and $h = 0.1$ (bottom panels).
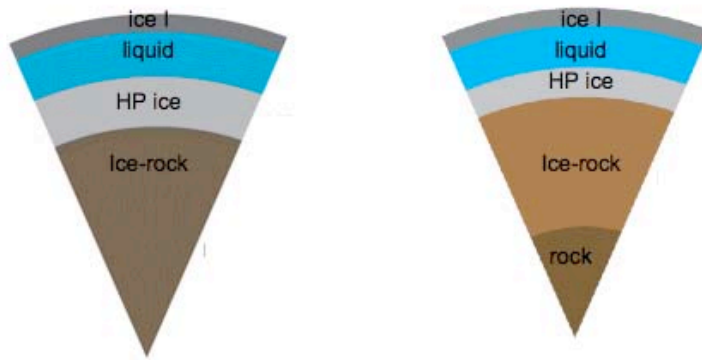
Figure 3. Two- and three-layer model of Titan. Titan's normalized moment of inertia as inferred by gravity data and hydrostatic equilibrium theory is $C/MR^2 = 0.34$ (*Iess et al.*, 2010). Titan's deep interior is likely composed of a mixture of ice and rock, indicating a partial differentiation. The outer layer is mainly composed of water in liquid and solid phase. The high-pressure (HP) polymorph ice that forms at the base of the outer layer has phases III, V, VI. At the center of Titan might be present a silicate or metallic core.

OBSERVATION AND MODELING OF ANTARCTIC ICE SHEET
AND OCEAN CIRCULATION INTERACTIONS

Eric Y. Larour (PI), Thermal and Cryogenics Engineering Section (354)
Dimitris Menemenlis, Climate, Ocean, and Solid Earth Science Section (324)
Eric Rignot, Department of Earth System Science, University of California Irvine
Todd Dupont, Department of Earth System Science, University of California Irvine
Michael Schodlok, JIFRESSE, University of California at Los Angeles

A. OBJECTIVES

Our objective was to couple the Ice Sheet System Model (ISSM), a new three-dimensional, high-resolution, higher-order model of the Antarctic ice sheet, with the global ocean, sea-ice, and ice shelf cavity model developed by the Estimating the Circulation and Climate of the Ocean, Phase II (ECCO2) project. The coupled model would then be integrated to provide a more realistic description of ice flow in Antarctica, ice shelf cavity melting, and ocean circulation around Antarctica, and to enable the development of sensitivity studies to determine the evolution of the Antarctic ice sheet into a warming climate.

B. APPROACH AND RESULTS

1. Coupling of ISSM and ECCO2

a. Anisotropic meshing routines to adapt to high-resolution ECCO2

ISSM [1] is a finite-element–based ice flow model, running on highly anisotropic meshes (see Figure 1, upper left frame), while ECCO2 is based on finite differences, running on regular grids (see Figure 2, upper right and lower left frames). In order to transfer data between ISSM and ECCO2 in a coupled run, two options are available. First, one can interpolate the data from the mesh onto the grid, and vice-versa. This solution is easy to implement on a technical level, but it introduces interpolation errors that can be detrimental to the coupling. Second, one can build the grids and meshes in such a way as to share common vertices. This approach, which was adopted in this project, is more difficult to implement on a technical level, but it has the advantage of introducing no interpolation error.

Figures 1 and 2 show meshes and grids for the Pine Island Glacier in Antarctica, in which the anisotropic mesh (on the grounded part of the ice sheet, Figure 1 upper left frame) was adapted to correspond exactly to a 1-km resolution grid on the ocean (Figure 2, upper left and lower right frames). The mesh was used to run ISSM in assimilation mode, in

order to infer basal drag (Figure 1 upper left frame), and improve best fit between modeled and observed velocities (Figure 1 lower left and lower right frames). The grid was used by ECCO2 to run an ocean circulation model to recover melting rates under the ice shelf.

b.   Data transfer routines between ISSM and ECCO2

In order to transfer data between ISSM and ECCO2, transfer routines were developed in Matlab. The routines extract from an ISSM mesh the data relevant to a specific ECCO2 grid, and expand this data onto the rest of the grid. Routines that do the exact opposite were developed, to transfer data from ECCO2 to ISSM.

These routines are simple enough that they could be encoded into different languages, such as C/C++/Fortran, used in ISSM and ECCO2.

2.   Test integrations, adjustment, and evaluation

a.   Downselection to Pine Island Glacier

Pine Island Glacier (Antarctica) has been experiencing strong acceleration in the last decade, which can only be explained by a sudden perturbation near the grounding line [2]. This perturbation is probably of oceanic origin [3], but the link between ocean warming and ice sheet acceleration has not been proven. As such, this basin presents a perfect testbed for implementing and evaluating a model that couples ice flow and ocean circulation. We therefore decided to focus the project on Pine Island Glacier itself.

The regional model implemented here relies on a 1-km grid, high-resolution ocean model as well as subkilometer resolution on the ice sheet. The component model configurations and the scenario for running the coupled ice/ocean model in Pine Island Glacier is explained in the next section.

b.   Ocean model component

The ocean model component of this study is integrated on a 640 by 640 horizontal grid with 1-km grid spacing and 50 vertical levels. Open boundary conditions are provided by a global ocean and sea ice solution generated by the Estimating the Circulation and Climate of the Ocean, Phase II (ECCO2) project [4]. Figure 4 shows an example result from this simulation, the Sea Surface Temperature (SST) on March 10, 1992. Of importance to realistic high-latitude ocean simulations is the correct representation of bottom boundary layers. Specifically, cold or saline dense water that sinks to the bottom needs to flow downslope in a thin (~20 m) layer. Because of limited horizontal and vertical resolution, this downslope current cannot be resolved explicitly in the above configuration. Instead, we have parameterized these currents using an embedded bottom boundary layer.

       c.   Transient runs using 0th order melting rates

At the time of this interim report, we were able to run a first-order projection of Pine Island ice flow dynamics, using constant melting rates provided by the ECCO2 ice shelf cavity model. The run lasted 20 years, starting from 2000. The velocity results are presented in Figure 3. Thicknesses provided by this run were then fed back to the ECCO2 model. Further work is still required to use computed melting rates from ECCO2, but this short-term projection of Pine Island Glacier was a good proof of concept, on which more complex and exhaustive models can be built.

3.   Short-term projection

This work was not carried out yet at the time of this interim report.

C.   SIGNIFICANCE OF RESULTS

1.   This is the first time an ice sheet model and an ocean circulation model are coupled. There is no example in the current literature of a coupled ice sheet/ocean circulation model. This project, once completed, should provide the first coupled model, to be used to model the sudden acceleration of outlet glaciers in Greenland and Antarctica. The link between ocean warming and glacier acceleration has never been modeled, which presents an opportunity for this project to yield some significant scientific results.

2.   Our effort will support ongoing NASA ocean and ice sheet missions (Jason-1, OSTM, QuikSCAT, AMSR-E, GRACE, and ICESAT-1) and international agency missions that provide data to NASA PIs (InSAR data from ERS, ALOS PALSAR, Radarsat-1, Envisat ASAR). This work will demonstrate the practical use of data assimilation techniques in coupled ocean and ice sheet modeling, thereby illustrating the usefulness of satellite data to produce better models; it also will provide guidance for the derivation of the science requirements of future missions such as GRACE-2, DESDYN-I, SWOT, and ICESAT-2.

D.   NEW TECHNOLOGY

No NTR was filed for new technology. The project resulted in new capabilities in ISSM and ECCO2 for transferring data between both models. The code is still in the validation/ verification stage, and should be completed in February 2011, at the end of this project. The project also resulted in the development of a bottom boundary layer parameterization for the ocean model used by the ECCO2 project: the Massachusetts Institute of Technology general circulation model (MITgcm). This parameterization is a critical component for realistic simulations of ocean/ice shelf interactions. The code is checked in the MITgcm CVS server so that it can be evaluated and improved by other MITgcm users.

E.   FINANCIAL STATUS

The total funding for this task was $200,000, of which $159,190 has been expended.

F.   ACKNOWLEDGEMENTS

G.   PUBLICATIONS

None.

H.   REFERENCES

[1]   M. Morlighem, E. Rignot, H. Seroussi, and E. Larour, "Spatial Patterns of Basal Drag Inferred Using Control Methods from a full-Stokes and Simpler Models for Pine Island Glacier, West Antarctica," *Geophys. Res. Lett., American Geophysical Union* **37** (2010): p. L14502.

[2]   A. J. Payne, A. Vieli, A. P. Shepherd, D. Wingham, and E. Rignot, "Recent Dramatic Thinning of Large West Antarctic Ice Stream Triggered by Oceans," *Geophys. Res. Lett., American Geophysical Union* **31** (2004) pp. 1–4.

[3]   A. J. Payne, P. R. Holland, A. P. Shepherd, I. C. Rutt, A. Jenkins, and I. Joughin, "Numerical Modeling of Ocean–Ice Interactions Under Pine Island Bay's Ice Shelf," *J. Geophys. Res. Oceans, American Geophysical Union,* **112** (2007): pp. 1–14.

[4]   D. Menemenlis, J. Campin, P. Heimbach, C. Hill, T. Lee, A. Nguyen, M. Schodlock, and H. Zhang, "ECCO2: High-Resolution Global Ocean and Sea Ice Data Synthesis," *Mercator Ocean Quarterly Newsletter* **31** (2008): pp. 13–21.

I.   FIGURES
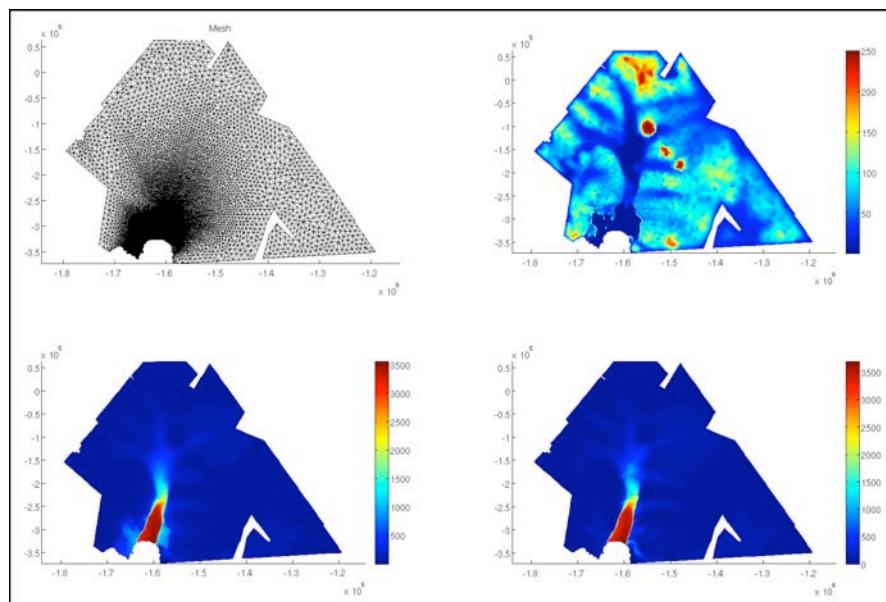
Figure 1. Model setup for Pine Island Glacier runs.
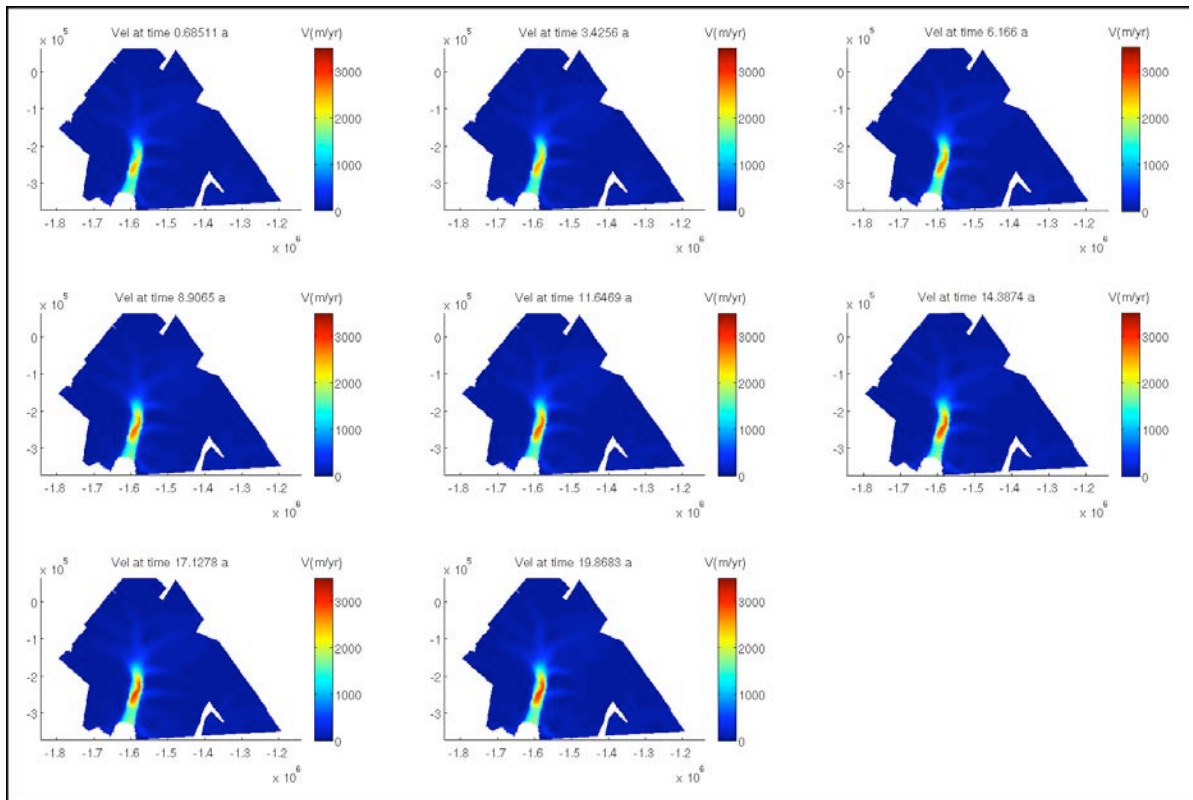
Figure 2. Data transfer between ice sheet and ocean circulation models.
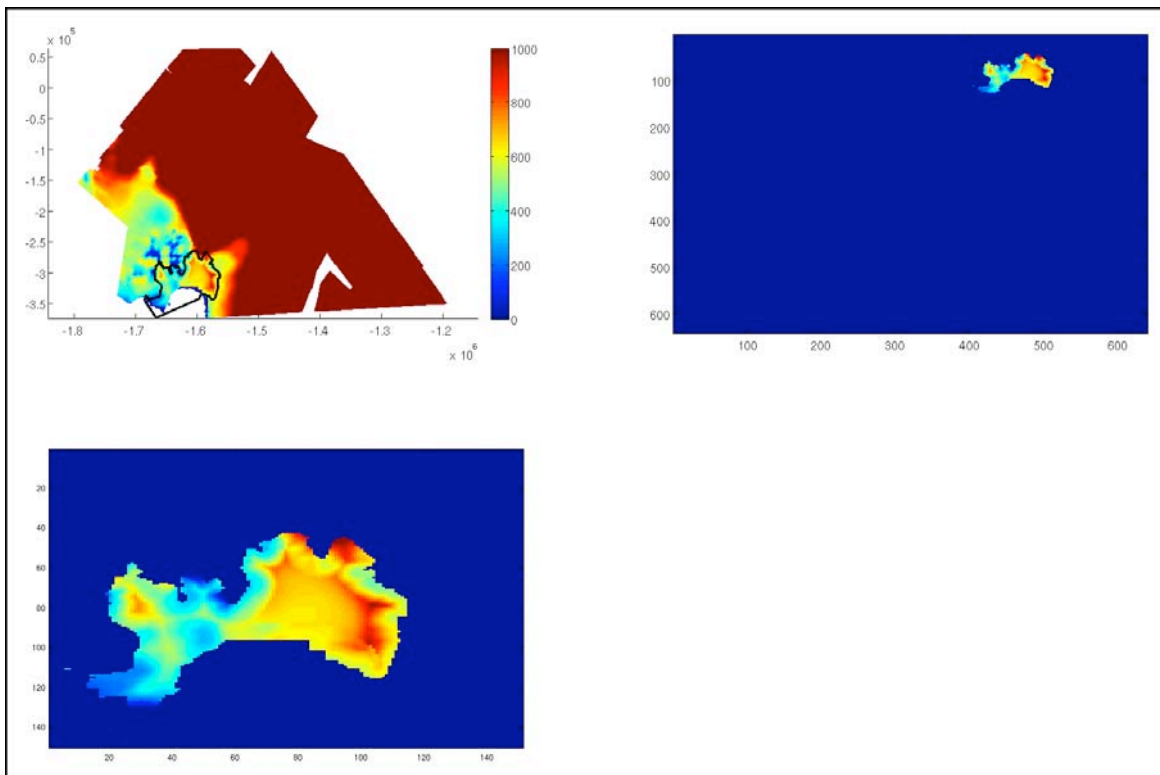
Figure 3. Transient velocities over 20 years projection, starting 1996, using constant melting rate provided by ECCO2 ice shelf cavity melting rate.
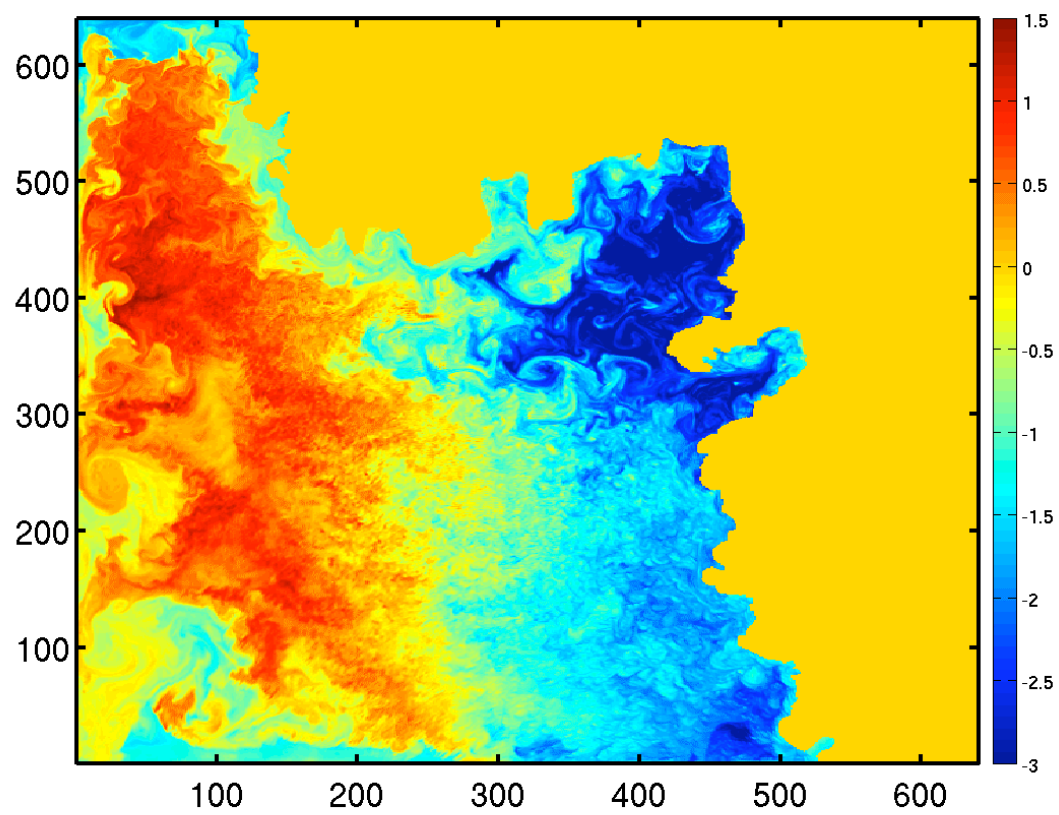
Figure 4. Sea surface temperature (10 March 1992).

PART 2: SURP ACTIVITIES

ASTRONOMY AND FUNDAMENTAL PHYSICS

AN ULTRA-LOW-POWER DIGITAL CORRELATOR
FOR A LUNAR RADIO INTERFEROMETER

Director's Research and Development Fund (DRDF)
Interim Report

JPL Task #1368

Charles R. Lawrence (PI), Astrophysics and Space Sciences Section (324)
Larry D'Addario, Tracking Systems and Applications Section (335)
Dayton L. Jones, Tracking Systems and Applications Section (335)
Gene Y. Lee, Systems Engineering Section (313)
Jacqueline N. Hewitt (Co-PI), Department of Physics, Massachusetts Institute of Technology
Sam Simmons, Department of Physics, Massachusetts Institute of Technology

A.   OBJECTIVES

As we observe objects in our universe at larger and larger distances, we see the universe as it was farther and farther back in time. The observations that probe the farthest back in the history of the universe are those of the cosmic microwave background (CMB), giving us information at a redshift of about $z=1000$. The most distant objects we have observed, other than the CMB, are quasars at a redshift of about $z=7$. A large part of our history, between redshifts 7 and 1000, remains unobserved. Probing this period in the history of the universe is at the frontier of cosmology today, and would allow us to explore the development of structure from the initial conditions shown by the CMB to the formation of the first stars and galaxies.

The most promising way to do this is to observe the redshifted 21-cm line of neutral hydrogen, the dominant constituent of the baryonic component of the matter in the universe at these redshifts. Ionospheric fluctuations and man-made interference will limit what can be done from the ground; the deepest integrations needed to fully probe the structures must be made from a place free from both. Thus, there is great interest in building a lunar radio array (LRA). The lunar farside has the advantages of negligible ionosphere and shielding from radio emissions from Earth. In addition, during part of its orbit, the Moon also provides shielding from low-frequency radiation from the Sun. The lunar surface thus provides a uniquely good environment for studying the Dark Ages and the EoR, even better than free space. Observation of neutral hydrogen is by far the most compelling astrophysical science proposed for the Moon.

Any scientifically interesting array will require a large field of view and a large collecting area in order to have sufficient survey speed to map the 21-cm structures. This drives the design to $\geq 10^4$ relatively small elements. All pairs of signals from the antenna elements have to be correlated; therefore, the computation burden of the correlator scales with the number of antenna elements squared. This scaling results in unacceptable power consumption if the technology of current Earth-based correlators is used. Our preliminary trade studies show that an LRA is feasible only if the power consumption of the correlator can be greatly reduced. Fortunately, this

seems to be possible. Over the last few years, the Earth science GeoSTAR project has led the development of low-power correlators for space. From the standpoint of sampling rate and multiplication rate per watt, the GeoSTAR 0.5V CMOS chips are excellent, and the technology is only improving with time, driven by commercial needs. However, there is more to a correlator than sampling and multiplication. In particular, the large number of interconnections between antennas and correlation elements is a major challenge. The goal of the proposed research is to develop a conceptual design for a correlator that will fall within the power (and by implication mass) envelope of a mission to the lunar farside.

B.   APPROACH AND RESULTS

We evaluated various design options for correlating a very large number of antenna elements, considering "station" size, beamformer type, and correlator architecture, and analyzing and optimizing the power used in each design. From this study several qualitative methods for minimizing power emerged:

1.   Arrange elements into a relatively small number of large "stations."
2.   Two-stage beamforming is generally the most efficient.
3.   Time domain fine delay beamforming is often more efficient than frequency domain fine delay.

We also investigated specific ways to reduce total power:

1.   Low-power CMOS technology
2.   Beamforming (one stage, two state, FFT)
3.   Matrix vs. pipeline architectures
4.   Scaling from existing correlator designs
5.   Study of published data sheets
6.   HP memory modeling tool for embedded ASIC memory: "CACTI"

We examined six different designs in detail for power usage, and found that the "LARC" concept for antenna design with two-stage beamforming uses the least power. In other words, a smaller number of stations with many elements per station is the most efficient design. Figure 1 shows the conceptual block diagram of the minimum power correlator, and Figure 2 shows the breakdown of power among various components.

We calculated that 534 stations with 85 elements and 16 beams per station, with 4,096,000 frequency channels of 150 MHz, could be correlated for 43 kW.

C.   SIGNIFICANCE OF RESULTS

We showed that a large lunar array could be correlated for an amount of power that is achievable on the far side of the Moon. The correlator, in other words, is not a showstopper for a large lunar array. As ASIC technology continues to develop, power levels can be expected to drop even further.

D.   FINANCIAL STATUS

The total funding for this task was $100,000, of which $79,518 has been expended.

E.   PUBLICATIONS

None.

F.   REFERENCES

None.
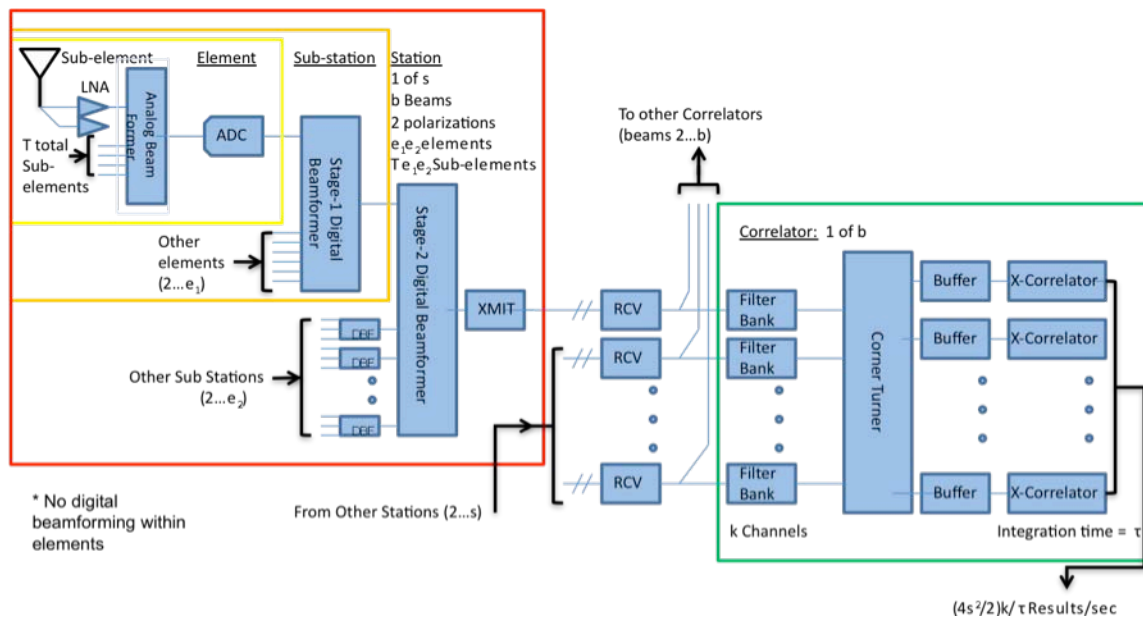
G.   FIGURES

Figure 1. Conceptual block diagram for a lunar radio array that would correlate 45,390 elements in 534 stations for a total power of about 43 kW.
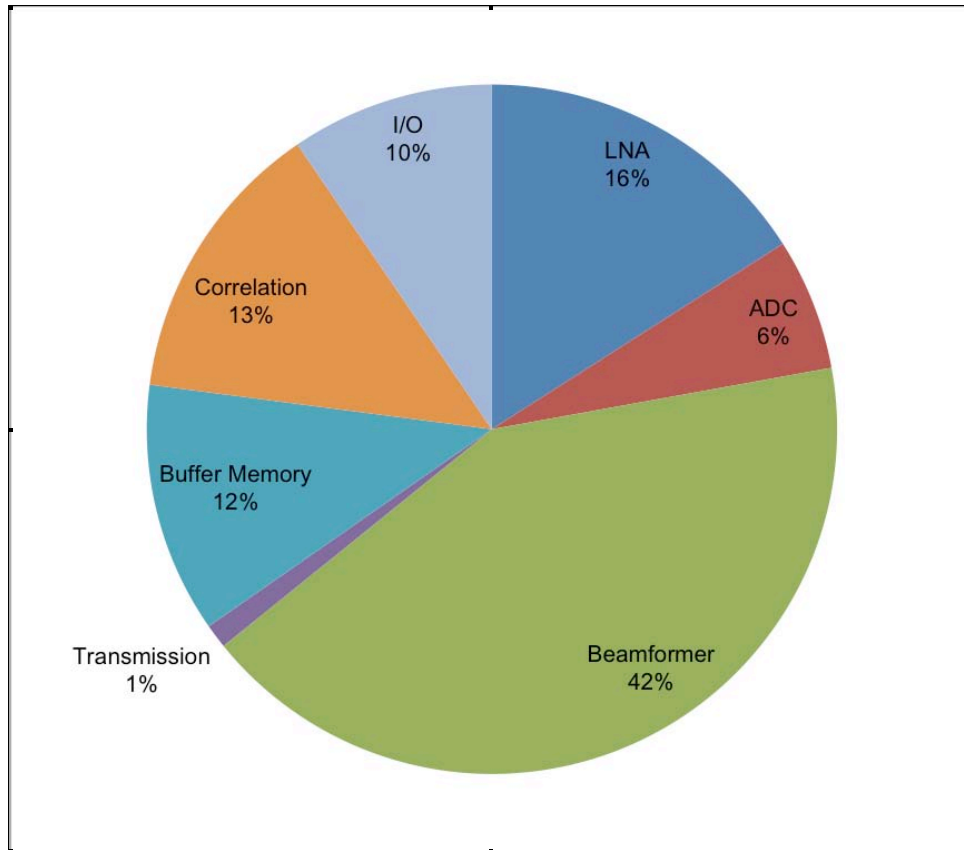
Figure 2. Power breakdown for the concept.

COMPUTATIONAL SCIENCE

LOW-POWER-CONSUMPTION LASERS FOR NEXT-GENERATION MINIATURE SPECTROMETERS FOR VENUS AND MARS

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1326

Siamak Forouhar (PI), Instrument Electronics and Sensors Section (389)
Alexander Soibel (Co-PI), Instrument Electronics and Sensors Section (389)
Lance E. Christensen (Co-PI), Atmospheric Observations Section (328)
Claire Gmachl (Co-PI), Department of Electrical Engineering, Princeton University

A.   OBJECTIVES

In situ detection of isotopic ratios of atmospheric trace gas species such as CO, CO_2, H_2O, H_2O_2, and NO have provided vital information about the formation and evolution of planetary bodies in our solar system. The presence and abundance of these chemicals are important indicators of geochemical processes, hydrothermal activity, and biological activity, if it exists. Measurements of these chemicals are of particular interest for the New Frontiers mission and Mars Exploration Program missions. The importance of trace gas detection was underscored by the selection of the tunable laser spectrometer (TLS) to be an instrument payload on the 2009 Mars Science Laboratory (MSL). Despite the remarkable instrument capabilities of TLS, the performance of next-generation tunable diode laser (TDL) spectrometers can be greatly improved by the development of mid-IR optical sources that require low electrical power to operate.

The objective of this task was to develop low-power quantum cascade (QC) semiconductor lasers operating in the 4.0–5.0-µm wavelength range. In the report, we detail progress toward our goal.

B.   APPROACH AND RESULTS

1.   Quantum design of 4.3-µm quantum cascade lasers

QC lasers were designed for emission near 4.3 µm (around CO_2 absorption) and 4.6 µm (around CO absorption) using the computational simulation tools developed at Princeton University. The design shown in Figure 1 as grown and fabricated emits near 4.6 µm, with performance characteristics shown in Figure 2.

2.   Growth of the lasers

Laser growth was accomplished by metal organic chemical vapor deposition through resources available to Princeton University. Laser devices were fabricated and tested at

Princeton University. Figure 2(a) shows light-current-voltage measurements for some of the best-performing devices measured at Princeton. Lasers were also delivered to JPL for further testing and analysis. JPL confirmed results obtained by Princeton. Figure 2(b) shows emission from a laser operating in continuous wave (CW) mode near room temperature as measured by JPL. The laser emission is around 4.6 μm, which is useful as it tunes over CO absorption lines.

3.   Quality control of facet coatings

High-quality laser facet coatings are an integral part of our strategy toward achieving low-input-power QC lasers. By increasing the reflectivity of both laser facets, we decrease threshold current, thereby reducing input power required for laser operation. We have initiated development for the capability to apply robust and uniform infrared facet coatings at JPL. We modified an electron beam deposition system in the JPL Microdevices Lab (MDL) dedicated to dielectric materials to add the ability to deposit on semiconductor laser bars and to improve the quality and reproducibility of the coatings produced by the system. Specifically, we completed fabrication of fixtures to properly secure and align fragile semiconductor material during the dielectric deposition process. We also augmented the system with the capability to precisely monitor atomic and molecular compositions present in the vacuum atmosphere through the addition of a mass spectrometer. This has allowed us to increase quality and reproducibility of facet coatings by alerting us when atomic and molecular species detrimental to facet coating quality are present within the deposition system.

4.   Characterization of the lasers

The ability to accurately measure performance characteristics is a crucial capability for a semiconductor. In this task, we assembled at JPL the necessary equipment to accomplish full characterization of QC lasers. This includes electrical and input/output power measurements with LIV (light-current-voltage) capability along with a Fourier transform infrared spectrometer (FTIR) for spectral characterization. We integrated and partially automated our equipment with the use of MATLAB programming and a control interface. The characterization setup assembled in this task is shown in Figure 3.

5.   Laser packaging

Quality packaging of semiconductor laser bars is a crucial capability for achieving high-performance devices. Packaging is meant to mechanically secure the laser die to a standardized mount convenient for end-user deployment. Moreover, by efficiently removing waste heat that is generated during laser operation, a high-quality packaging process allows the laser device to achieve maximum operating temperature. We developed a complete packaging processs — including dicing/cleaving of processed wafers, attaching laser die to mounts/submounts, and wire bonding electrical contacts — for our QC laser devices.

We designed our die attach process to provide a high degree of mechanical integrity while minimizing thermal resistance. Using a $Au_{10}Sn_{90}$ fluxless solder preform, we achieved high-quality die attach to C-mount and TO-9 standardized packaging systems, as shown in Figures 4(a)–4(c). Figure 4(d) shows a bonding process that has yielded excellent facet

alignment with the edge of mounting surface and visual evidence of near 100% adhesion of the die to the attach medium. Destructive die shear tests, the results of which are shown in Figure 4(e) for multiple devices, indicate shear strength that exceeds Mil. Std. 883G (600 g force for a 1×1 mm die). Our current packaging process is finalized with wire bonding the laser die to the submount. Figures 4(a)–4(d) show multiple wire bonds from the laser die top contact to the laser package. The wire bonds are generally high quality, as most surpass Mil. Std. 883G with >3 g of force required for failure. The results of a series of destructive bond pull tests are shown in Figure 4(f).

C.   SIGNIFICANCE OF RESULTS

The primary objective of this task was to develop low-input-power QC lasers operating in the 4–5-μm wavelength range. The completed tasks and accomplishments outlined above are progress toward that goal. We demonstrated room-temperature CW operation of QC lasers operating at 4.6 μm, a wavelength useful for high-sensitivity CO detection. We also developed a foundation for fabricating high-performance QC lasers at JPL with the development of an advanced packaging process. With the test and characterization facility developed within this program, JPL is prepared for QC laser fabrication that can impact future space missions.

Measurements of chemical isotopes are just one area where this work will impact JPL missions and programs. Two methods are available for performing in situ measurements of isotopic ratios of atmospheric trace gases: isotope ratio mass spectrometry (IRMS) and tunable laser spectroscopy (TLS). The latter offers a much simpler method for isotope analysis since tunable laser spectrometers can be made much more compact than traditional IRMS instruments and chemical conversion is not required (e.g., IRMS requires methane to be oxidized to CO_2). However, a fully robust and capable TLS system is not feasible if the laser source requires multi-watt input power for operation. Through the work done within this DRDF research, we have advanced our ability to produce low-input-power, mid-infrared laser sources.

D.   FINANCIAL STATUS

The total funding for this task was $100,000, all of which has been expended.

E.   ACKNOWLEDGEMENTS

Collaborators on this project included: Richard Cendejas, Department of Electrical Engineering, Princeton University; Matthew D. Escarra, Department of Electrical Engineering, Princeton University; Loan Le, Department of Electrical Engineering, Princeton University; Arjun Vijayakumar, Department of Electrical Engineering, Princeton University; Kale J. Franz, Instrument Electronics and Sensors Section (389), and (formerly) Department of Electrical Engineering, Princeton University; Clifford Frez, Instrument Electronics and Sensors Section (389); and Yueming Qiu, Instrument Electronics and Sensors Section (389).

F.   PUBLICATIONS

None.

G.   REFERENCES

None.

H.   APPENDIX: Noteworthy Ancillary Results of Princeton–JPL Collaboration

As a result of this collaboration between Princeton University and JPL, JPL was able to hire Kale Franz, a graduate of Prof. Claire Gmachl's research group at Princeton University. Dr. Franz brings with him to JPL advanced knowledge of QC laser operation, design, and fabrication. His skills augment JPL's ability to deliver a full spectrum of semiconductor laser devices for advanced space applications.

I.   FIGURES
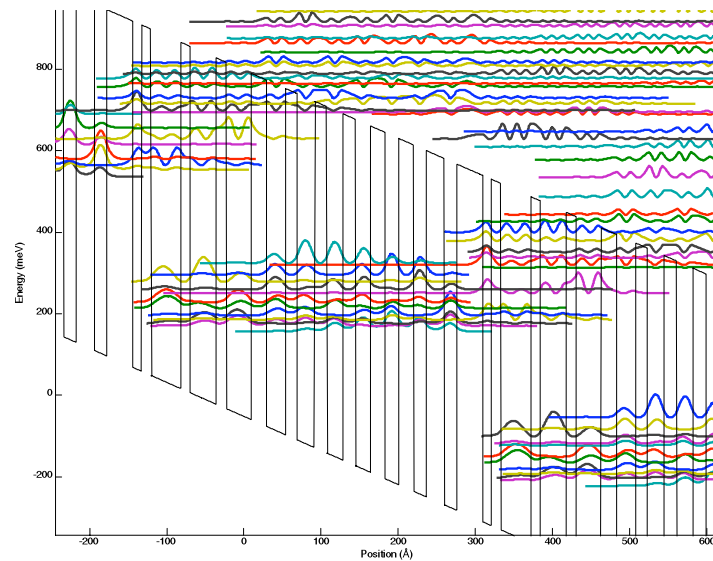
Figure 1. A QC laser structure designed and grown by Princeton University. As shown in the following figures, it is capable of room-temperature CW emission.
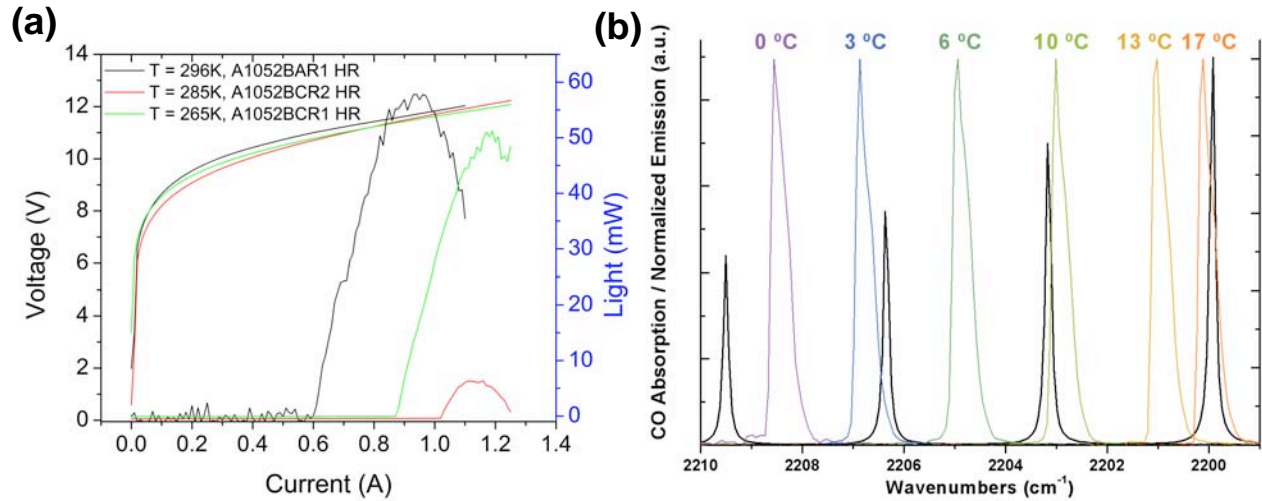
Figure 2. (a) Top performance achieved from several different laser ridges of the QC laser structure shown in Figure 1. The curves show voltage vs. current and light vs. current for three different lasers operating in CW mode. (b) Normalized emission of a QC laser operating in CW mode near room temperature. The laser operates above room temperature; however, this temperature region was chosen to show the laser's thermal wavelength turning over several CO absorption peaks, shown in the above figure with a bold black line.
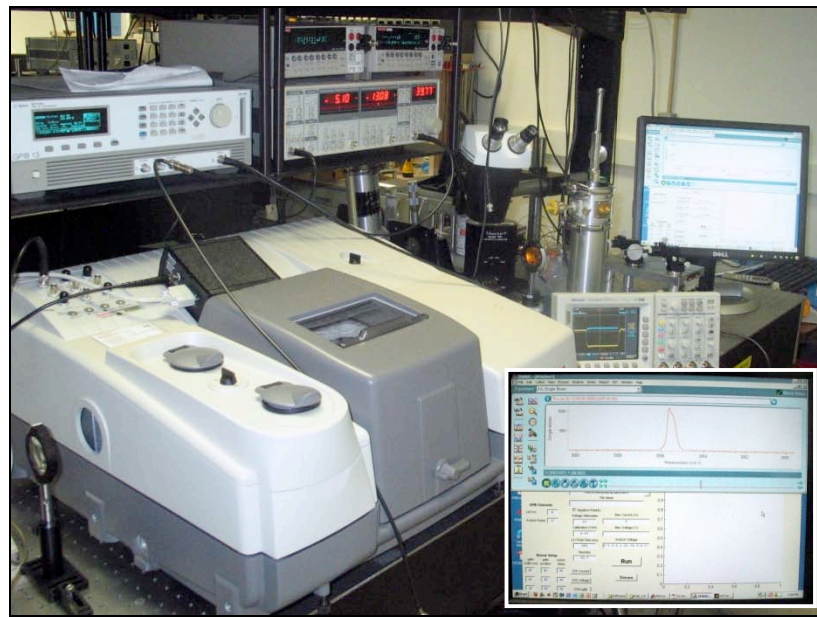
Figure 3. The QC laser characterization setup assembled for characterization work at JPL. The system includes an FTIR, power sources, and temperature control, and is controlled through a customized MATLAB interface.
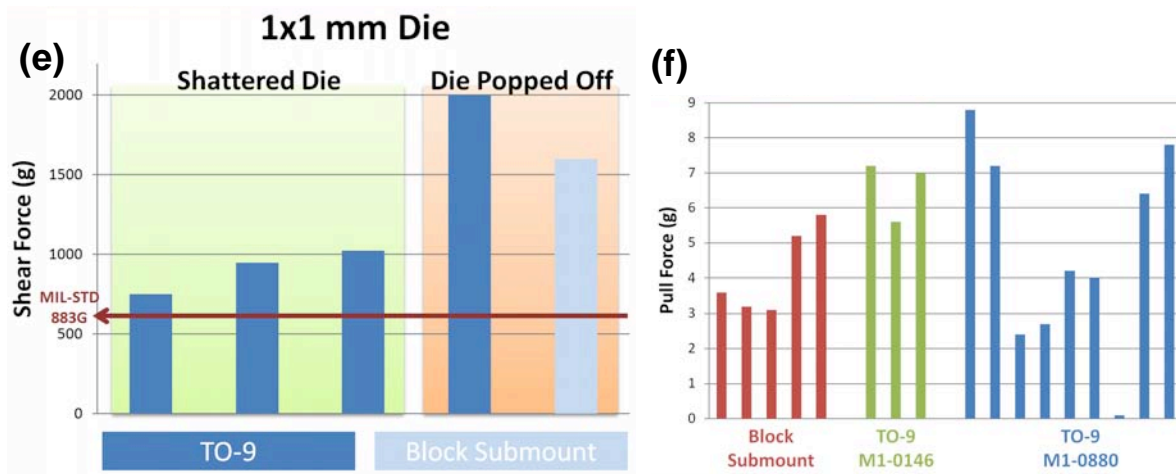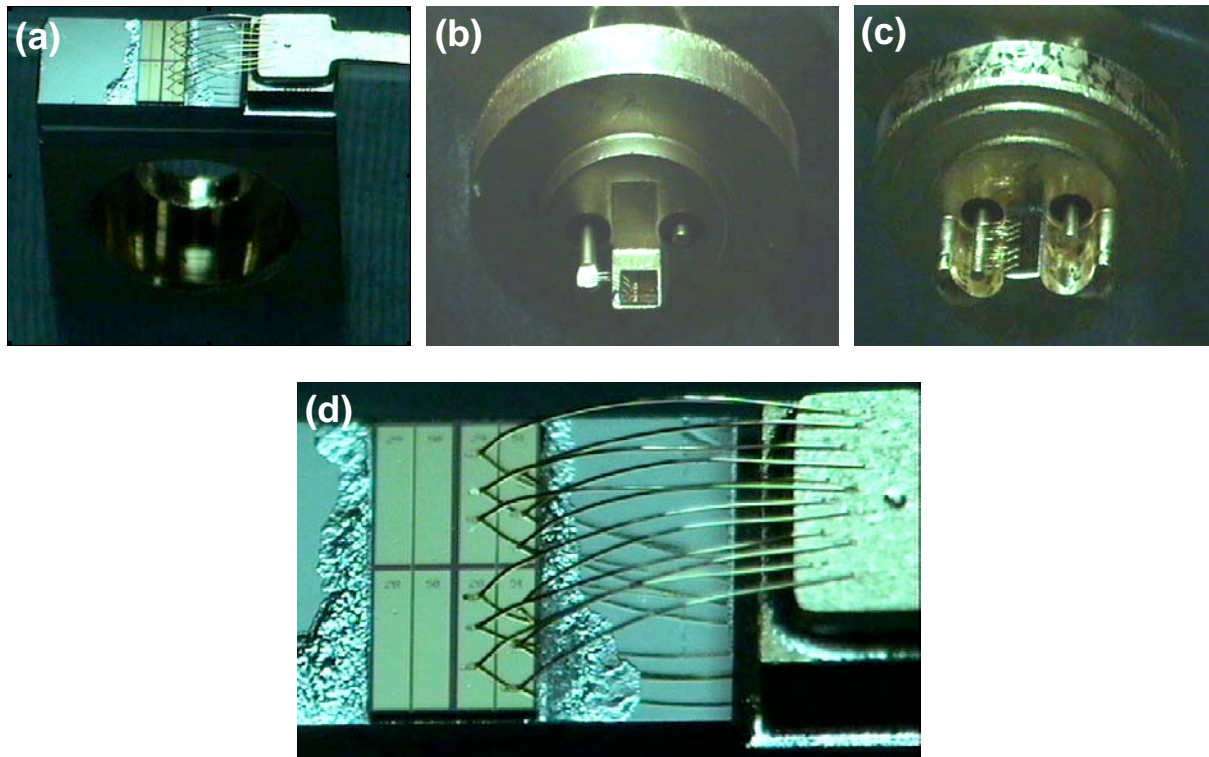
Figure 4. Packaging of laser die developed by JPL through this work. Panels (a)–(c) show laser die attached to a C-mount and two different TO-9 packages. Panel (d) shows in detail a high-quality die attach along with wire bonding of the laser top contact. Panels (e) and (f) show the results of destructive shear force and wire pull tests for several devices, respectively. The tests indicate a high-quality packaging process that passes Mil. Std. 883G.

DEPLOYABLE STRUCTURES

TESTING GENERAL RELATIVITY ON COSMOLOGICAL SCALES WITH WEAK GRAVITATIONAL LENSING

Jason D. Rhodes (PI), Astrophysics and Space Sciences Section (326)
Robert R. Caldwell, Department of Physics and Astronomy, Dartmouth College
R. Ali Vanderveld, Astrophysics and Space Sciences Section (326)

A.   OBJECTIVES

Our objective was to develop new tests of gravitational physics on cosmological scales. We focused on weak gravitational lensing observables, as the Dark Energy Task Force [1] has recommended lensing of galaxies and clusters as a promising method for constraining dark energy and modified gravity (possible modifications to General Relativity that would explain current cosmological measurements). This is because the dark matter power spectrum derived via weak lensing is sensitive to the effects of the expansion history of the Universe. We further expect that modifications of General Relativity will affect other weak lensing observables, since such modifications affect both how matter gravitates, and thus how large-scale structures form, and how photons move in gravitational potentials. Gravitational lensing observations are therefore the ideal probe of modified gravity, as they depend on the phenomenology of both structure formation and light bending.

One new way in which weak lensing can be used to constrain gravitational physics is in the nonlinear regime. The second-order correction to the lensing deflection angle contains new information, such as the degree of nonlinearity in gravitational superposition, which may provide clues to a possible departure from General Relativity. Although this correction contributes negligibly to the dark matter power spectrum (the two-point correlation function of density fluctuations), it may be accessible through the dark matter bispectrum (the three-point correlation function). Our goal was to determine how nonlinear weak lensing statistics can be used to constrain deviations from General Relativity in a model-independent way.

B.   APPROACH AND RESULTS

We carried out calculations within a parameterized-post-Friedmannian (PPF) description of gravitation. The PPF formalism is a framework for constraining generic deviations from General Relativity in the cosmological weak field (i.e., small gravitational potential) regime. This framework is based on the very successful parameterized-post-Newtonian (PPN) framework, which expresses the nonlinear equations of theories of gravitation in terms of a post-Newtonian expansion, wherein it is assumed that the deviations from Newtonian theory are small. This assumption will be true in the limit of small velocities and gravitational potentials.

Weak lensing observables fall within these limits.

Moreover, assuming metric theories where the Einstein equivalence principle is still valid, this framework parameterizes any allowable deviations from General Relativity in terms of a number of PPN parameters. These parameterize various possible characteristics of theories of gravitation [2], such as how much space curvature is produced by a unit rest mass, and how much nonlinearity is in the superposition law for gravitation. There is an associated parameter for each of these characteristics. The weak field regime of almost any theory of gravitation will have certain values for each. Therefore, it is the cosmological analogs of these parameters, in the PPF framework, that will be constrained with lensing data. Initially, these parameters will be used as a null test of General Relativity, as we look for any clues as to whether Einstein's theory is valid on the largest scales or if novel gravitational effects are responsible for the cosmic acceleration. Later, these results can be translated into constraints on specific modified gravity models.

We computed light deflection within this PPF framework, to second order in the Newtonian gravitational potential. From this, we calculated image distortions as a function of these parameters. Performing these complicated calculations was the primary goal of our proposal. Having done this, we have laid the groundwork for the next stage of this ongoing research (beyond the scope of this SURP project). The next step will be to compute the standard weak lensing statistics, including the weak lensing power spectrum and bispectrum. We can then use combinations of these statistics to constrain PPF parameters and hence the allowable deviations from General Relativity on cosmological scales. For example, the weak lensing bispectrum vanishes in linear theory and thus only depends on the nonlinear gravitational potential. This observable will therefore depend on the PPF parameter that controls how much nonlinearity is in the superposition law for gravity. A constraint on this parameter would provide valuable information regarding what modifications of General Relativity are allowable on large scales. As proof of concept of our method, and in order to determine the size of the deviations to be expected, we also determined these PPF parameters in a leading modified gravity theory — f(R) gravity [3].

C.   SIGNIFICANCE OF RESULTS

We achieved our primary objective — to develop a new test of General Relativity on cosmological scales — by utilizing nonlinear weak lensing statistics. In the future, this will enable us to determine how well these new tests will perform with current and future data sets. This new test of gravity will enhance the science output of future weak lensing surveys such as the NASA/DOE Joint Dark Energy Mission (JDEM) and the ESA Cosmic Visions Euclid mission.

D.   FINANCIAL STATUS

The total funding for this task was $100,000, all of which has been expended.

E. PUBLICATIONS AND PRESENTATIONS

[A] Robert R. Caldwell and R. Ali Vanderveld, "Second-Order Weak Lensing as a Test of New Gravitational Phenomena," in preparation.

[B] R. Ali Vanderveld, "Testing Gravity on Cosmological Scales with Weak Lensing," Cosmo 09, Geneva, Switzerland, September 7, 2009.

[C] R. Ali Vanderveld, "Testing Gravity on Cosmological Scales with Weak Lensing," September 29, 2009, UC Berkeley; October 16, 2009, University of Chicago; October 19, 2009, Fermilab; October 27, 2009, Perimeter Institute; October 29, 2009, Canadian Institute for Theoretical Astrophysics; November 15, 2009, Caltech.

F. REFERENCES

[1] Andreas Albrecht, et al., "Report of the Dark Energy Task Force," arXiv:astro-ph/0609591, 2006.

[2] Cliff Will, "Theory and Experiment in Gravitational Physics," Cambridge University Press, Cambridge, UK, 1993.

[3] Sean Carroll, Vikram Duvvuri, Mark Trodden, and Michael Turner, "Is Cosmic Speed-up Due to New Gravitational Physics?" Physical Review D 70 (August 19, 2004): p. 043528.

DEVELOPMENT OF A 150-GHz MMIC MODULE FOR MEASUREMENTS OF THE POLARIZATION OF THE COSMIC MICROWAVE BACKGROUND RADIATION

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1312

Lorene A. Samoska (PI), Instrument Electronics and Sensors Section (389)
Todd C. Gaier, Instrument Systems Implementation and Concepts Section (382)
Pekka Kangaslahti, Instrument Systems Implementation and Concepts Section (382)
Mary Soria, Instrument Systems Implementation and Concepts Section (382)
Heather Owen, Instrument Systems Implementation and Concepts Section (382)
Sarah E. Church (Co-PI), Kavli Institute for Particle Astrophysics and Cosmology,
and Department of Physics, Stanford University
Patricia Voll, Kavli Institute for Particle Astrophysics and Cosmology,
and Department of Physics, Stanford University
Judy M. Lau, Kavli Institute for Particle Astrophysics and Cosmology,
and Department of Physics, Stanford University
Matt M. Sieth, Kavli Institute for Particle Astrophysics and Cosmology,
and Department of Physics, Stanford University
Dan van Winkle, Kavli Institute for Particle Astrophysics and Cosmology, Stanford University

A.   OBJECTIVES

The objective of this work was to design, fabricate, and test a heterodyne amplifier module intended to operate over the frequency range 140–170 GHz. The module is based on Monolithic Millimeter-Wave Integrated Circuit (MMIC) Low-Noise Amplifiers (LNAs). It was designed to be scalable in order to allow the development of arrays with hundreds of elements for measuring the polarization of the Cosmic Microwave Background (CMB). A space mission based on such arrays could be used to search for the signature of inflation, and other fundamental physical processes, in the curl modes (B-modes) of the polarization of the CMB.

In the last few decades, experiments such as the Cosmic Background Imager (CBI), Degree Angular Scale Interferometer (DASI), and WMAP satellite have provided state-of-the-art CMB polarization measurements over the frequency range of 5–100 GHz, using HEMT amplifiers that offer low noise, low power dissipation, high reliability, and inherently wide bandwidths. However, these experiments were equipped with a small number of radio receivers painstakingly assembled by hand, and a significant increase in capability is needed to detect the much fainter B-mode component of the CMB. Recent advances in low-noise MMIC amplifiers, coupled with industry-driven advances in high frequency signal interconnects and in the miniaturization and integration of many standard components, have decreased the technological gap that must be overcome to create radio cameras with hundreds of pixels for operation up to 200 GHz and that are sensitive to a wide (20–25%) simultaneous bandwidth.

The goal of this work was to deliver a module that is demonstrated to operate at cryogenic temperatures with good noise performance and wide bandwidth and that is fabricated using techniques that will enable the mass production of very large focal planes intended for radiometry, interferometry, and spectroscopy. This technological demonstration builds on the successful development by JPL of 40-GHz and 90-GHz correlation radiometers for the QU Imaging ExperimenT (QUIET) that is currently making observations of CMB polarization from the Atacama Plateau, and 90-GHz heterodyne modules for the MAS (Millimeter Array Spectrograph) experiment, which is under development. This work will extend the capability of this technology to 150 GHz, an important frequency for separating the CMB polarization signal from astrophysical foregrounds, and for other important astrophysical measurements such as studying the hot gas in galaxy clusters via the Sunyaev-Zel'dovich effect.

B. APPROACH AND RESULTS

1. <u>Design</u>. A block diagram of the module design is depicted in Figure 1. Incident radiation is collected by an external feedhorn antenna and directed through a WR-5 waveguide cavity to the components in the block diagram, where the RF signal is subsequently amplified by a series of low-noise amplifiers (LNAs) and downconverted to intermediate frequency (IF) by a subharmonic mixer. The local oscillator (LO) signal for the mixer is fed into the module housing via WR-10 waveguide. E-plane probes extending into the waveguide cavity provide the transitions from the WR-10 and WR-5 waveguide inputs to the planar MMIC components. The LNAs, mixer, and E-plane probes populating the module were selected based on availability and the desired frequency band (140–170 GHz).

The LNAs in the module amplify the incoming RF signal and pass it to the mixer for downconversion. The LNA design is based on HEMT technology and is described in detail in [1]. The module will operate at 30 K, where the minimum noise temperature of the LNA is less than 100 K. The assembled module is shown in Figure 2.

2. <u>Expected Performance</u>. The expected noise performance of the module design can be roughly estimated from the gain and noise temperature data for the various module components. This calculation indicates that at room temperature the module should have a noise temperature of approximately 450 K. When the physical temperature of the module is 30 K, a noise temperature of approximately 125 K is expected (Figure 3).

3. <u>Measured Results</u>. We attempted testing of the module as a receiver using a WR-5 waveguide horn antenna and a Gunn oscillator to provide the LO signal (Figure 4). At this time, we have observed the local oscillator signal leaking through to the output of the last LNA. It has altered the LNA bias condition, and may be causing substantial gain compression. We have been able to measure a receiver noise temperature of 2100 K at 160 GHz. This noise temperature is much greater than anticipated and is most likely due to the compression in the second LNA from the local oscillator signal. A bandpass filter to reject W-band has been designed and is in fabrication, and will suppress the LO leakage onto the LNA. The filter design is shown in Figure 5, and was designed for a 3-mil alumina substrate. The filter will fit into the module cavity between the second LNA and the mixer. We simulated the S-parameters of the filter and

they are shown in Figure 6. The rejection at W-band below 85 GHz is greater than 30 dB. The mixer chip has been measured on-wafer to have approximately 20 dB of LO to RF isolation. The addition of the filter will increase this by 30 dB.

C.   SIGNIFICANCE OF RESULTS

The prototype module that has been designed under this SURP program will be used to develop multiple instruments for astrophysics radio telescopes, both on the ground and in space. These include the Chajnantor Inflation Probe (CHIP), Space Heterodyne Inflation Probe (SHIP), CMB Pol, SAFIR, CALISTO, and future missions requiring compact, scalable, low-power heterodyne receivers and receiver arrays.

D.   NEW TECHNOLOGY

We reported the completed module development in NTR #47664, "Development of a 150-GHz MMIC module prototype for large-scale CMB radiation experiments."

E.   FINANCIAL STATUS

The total funding for this task was $100,000, all of which has been expended.

F.   ACKNOWLEDGEMENTS

We would like to acknowledge Douglas Dawson of JPL for assistance with the bandpass filter design.

G.   PUBLICATIONS AND PRESENTATIONS

[A]   Patricia Voll, Judy M. Lau, Matt Sieth, Sarah E. Church, Lorene A. Samoska, Pekka Kangaslahti, Todd C. Gaier, Dan Van Winkle, and Sami Tantawi, "Development of a 150-GHz MMIC Module Prototype for Large-Scale CMB Radiation Experiments," SPIE Astronomical Telescopes and Instrumentation, San Diego, California, June 26, 2010.

H.   REFERENCES

[1]   Pekka Kangaslahti, David Pukala, Todd Gaier, William Deal, Xioabing Mei, and Richard Lai, "Low-Noise Amplifier for 180-GHz Frequency Band," *IEEE Int. Microwave Symp. Dig.*, June 15–20, 2008, Atlanta, Georgia: pp. 451–454.
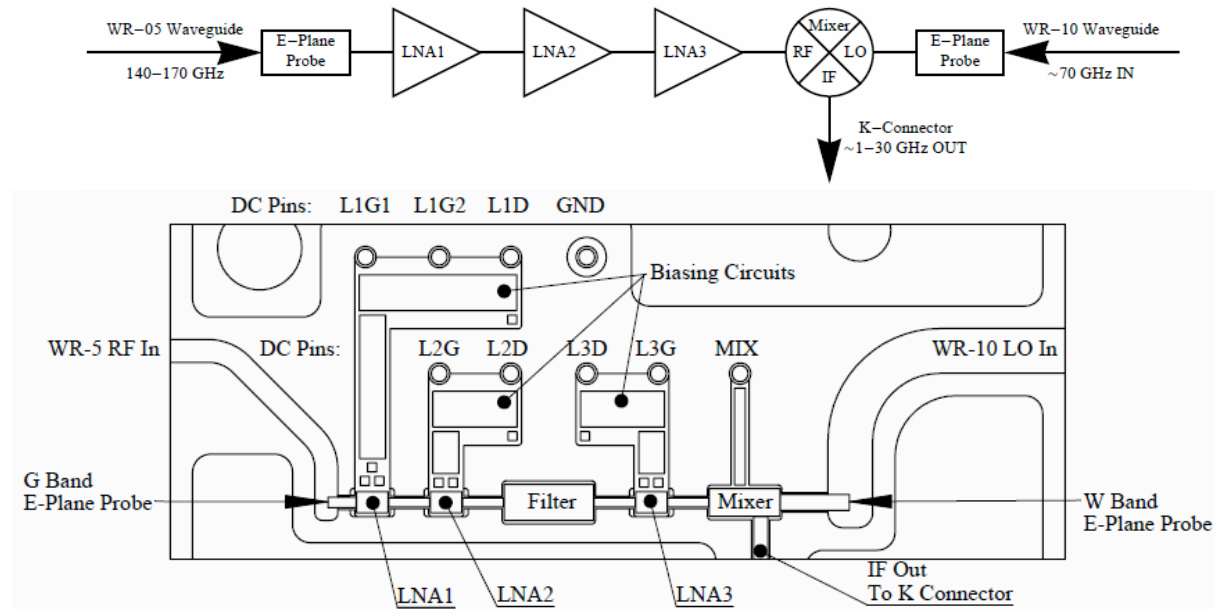
Figure 1. Top: A block diagram of the components within the module. Bottom: Sketch showing layout of the chips and other components.
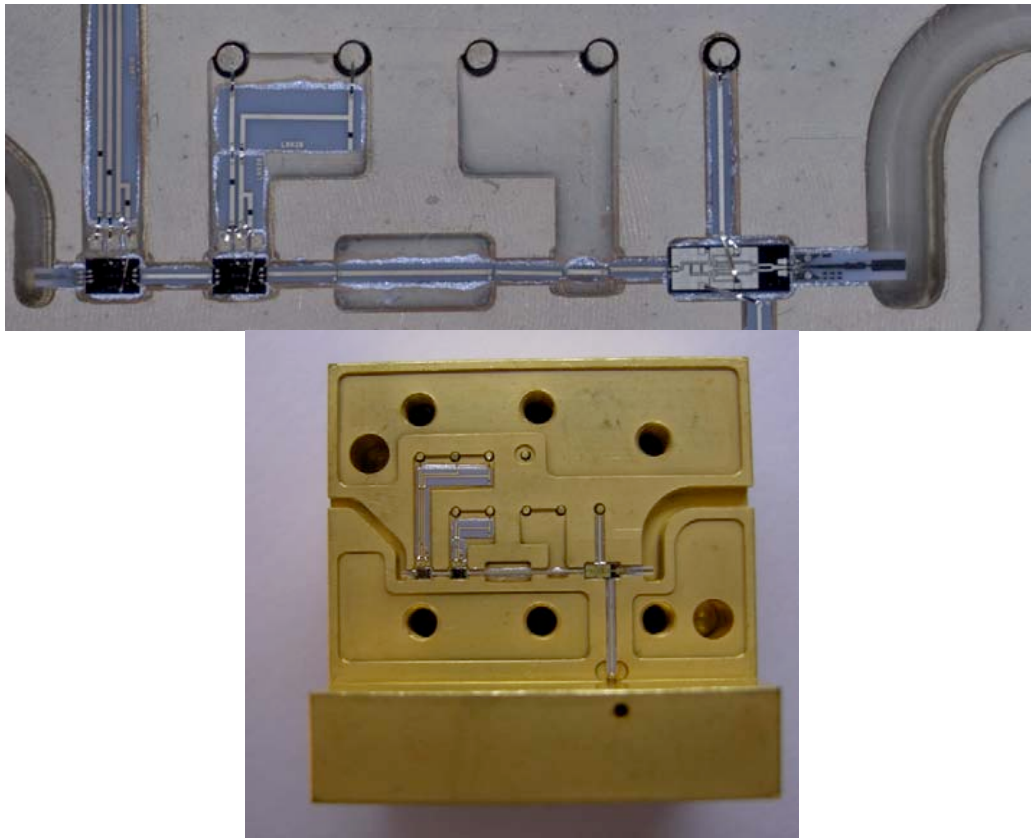
Figure 2. 150-GHz MMIC receiver module, assembled with lid open.
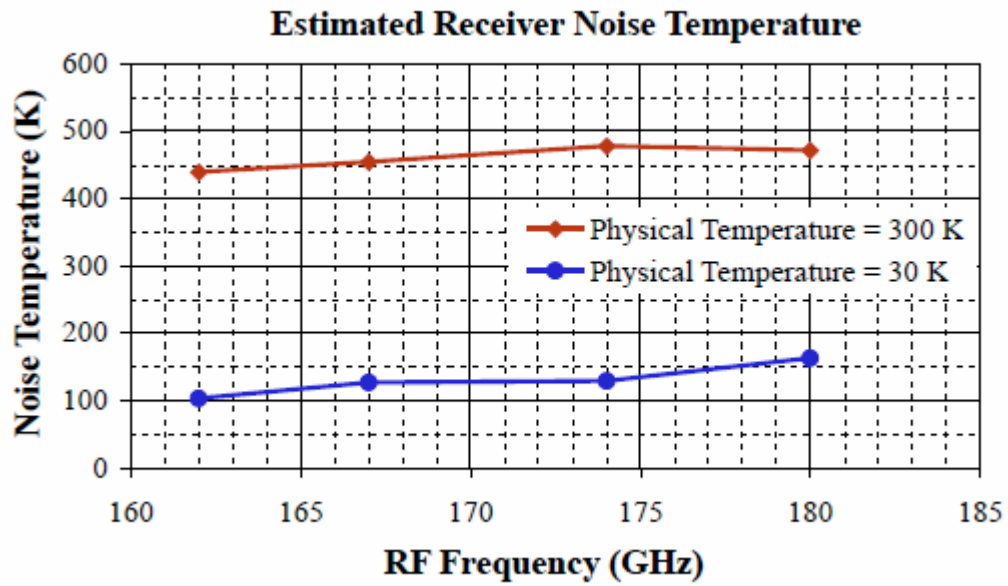
Figure 3. Estimated noise temperature of the module at 30 K and 300 K physical temperature. This estimate is based on the data for the individual components used in the module.
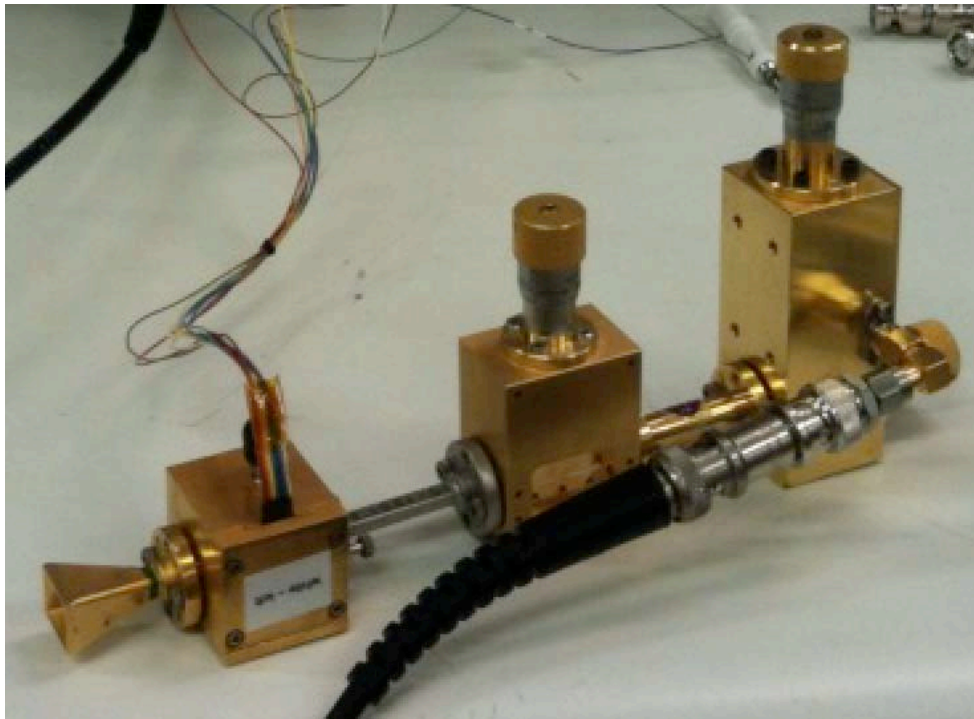
Figure 4. 150-GHz MMIC receiver module, with local oscillator applied.
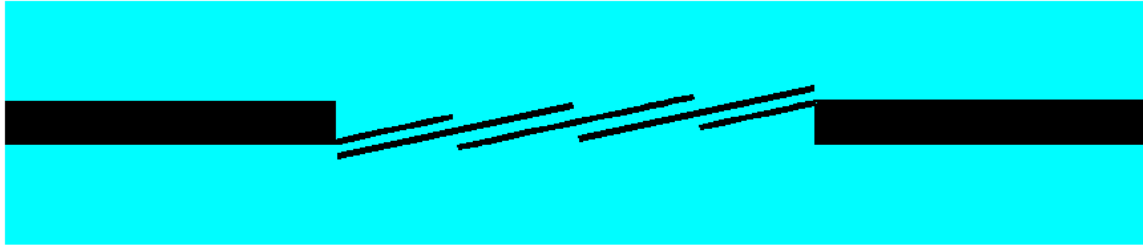
Figure 5. Autocad drawing of parallel coupled-line bandpass filter for 120-170 GHz, designed for a 3-mil alumina substrate.
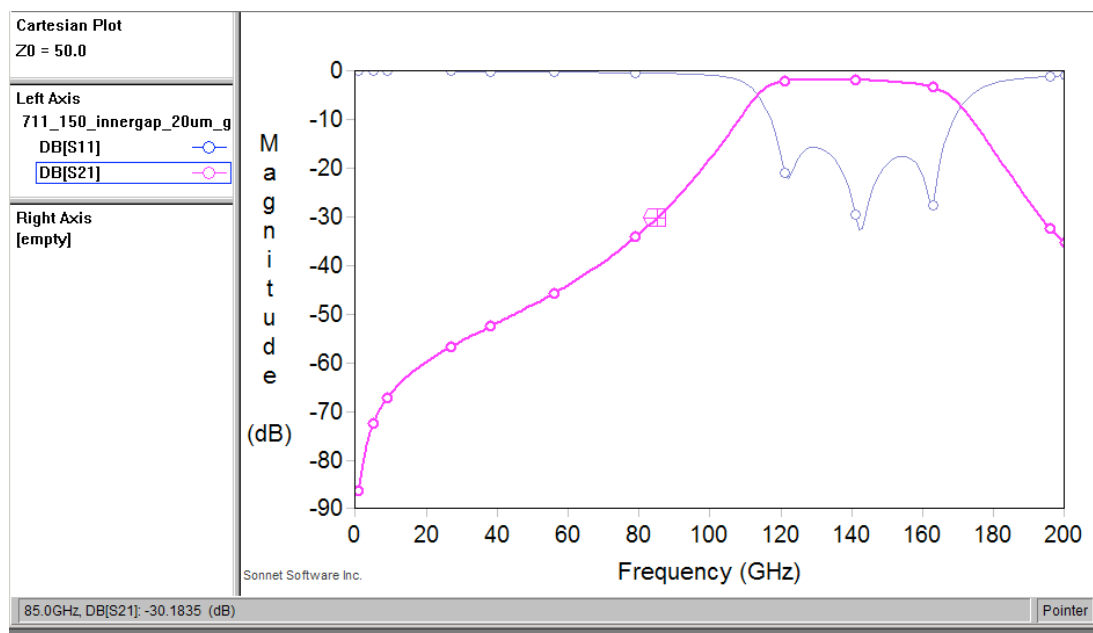
Figure 6. Simulated S-parameters of the bandpass filter, showing greater than 30 dB of rejection below 85 GHz.

A 1900-GHz LOCAL OSCILLATOR CHAIN FOR EXPANDED MULTI-PIXEL GALACTIC SURVEY ON THE STRATOSPHERIC THz OBSERVATORY

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1313

Imran Mehdi (PI), Instrument Electronics and Sensors Section (389)
Chris Walker, Department of Astronomy, University of Arizona
Choonsup Lee, Instrument Electronics and Sensors Section (389)

A.   OBJECTIVES

The objective is to design the next generation (post-HIFI) of a local oscillator (LO) source centered at 1900 GHz. This next-generation source is the enabling technology that would allow the potential to place an expanded, multi-pixel hot electron bolometer (HEB)–based heterodyne receiver system on an already funded long-duration balloon platform, Stratospheric Terahertz Observatory (STO). STO is a collaborative project among the University of Arizona (Walker-PI), JPL, Caltech, JHAPL, and several other institutions. The STO mission is funded by NASA's Astrophysics Research and Analysis (APRA) program. The initial observational goal of STO is to make high spectral (<0.1 km/s) and angular resolution (~1') maps of the Galactic Plane in the astrophysically important transitions of C+ at 1.9 THz and N+ at 1.46 THz. To achieve the angular resolution requirement, STO will have an aperture of ~80 cm. To achieve the target spectral resolution, STO will utilize a heterodyne receiver system. The proposed LO development is the key element in determining the size of the receiver array. While the source developed for HIFI is sufficient to pump two HEB channels, the next generation source will deliver more than 10 microwatts of output power that should be sufficient to pump 4-8 HEB channels. The proposed source with its expanded receiver system could serve as the testbed for a THz instrument for the next round of MIDEX proposals.

B.   APPROACH AND RESULTS

To increase the output power at 1900 GHz, a number of steps have been taken. The approach that will be followed is shown in Figure 1. The proposed 1.9-THz LO source will be implemented as a ×2×3×3 chain. We will design a 4-chip 200-GHz doubler that will provide 80 mW of output power. This will be used to pump a 2-chip 600-GHz tripler. It is expected that such a tripler will provide more than 5 mW of output power. We predict that this will be enough to pump the 1.9-THz tripler to produce more than 10 microwatts of output power. If possible, the chain can be cooled to 150 K to meet this goal. A 4-chip power amplifier module will be built using the approach shown in Figure 2. This block has already been built. Output power is expected to be in the 400 mW range around 106 GHz.

We recently demonstrated the use of multiple multiplier chips in a single waveguide block to increase output power. The 4-chip 200-GHz doubler design has been completed and blocks are being fabricated at the University of Arizona. This block will utilize four membrane-based doubler chips. The input signal is divided equally between four chips, whereas the output signal from the four chips is combined via a Y power combiner. Output power of 80 mW at 212 GHz is around a factor of 4 higher from what can be achieved with a single-chip membrane-based implementation. Each chip is expected to have around 20% efficiency, which will provide an ~80-mW signal at 212 GHz.

Power from the doubler will be used to pump a tripler working in the 600-GHz range. Results from a single-chip implementation are shown in Figure 3. Two blocks (SN3 and SN8) have been tested. A single block that uses two of these chips will be used to produce more than 5 mW to drive the final stage tripler. The 1900-GHz tripler chip has been designed to produce around 10 microwatts and is shown in Figure 4.

C.   SIGNIFICANCE OF RESULTS

The C+ line (at ~158 microns) is the strongest Galactic cooling line and will be the premier diagnostic tool for studying the life cycle of the interstellar medium (ISM) in the Milky Way and other nearby galaxies. The STO C+ survey, together with pointed observations from Herschel, will provide the "Rosetta Stone" for translating the global properties of distant galaxies into reliable estimators of star formation rate and state of the ISM. The observing strategy for the survey is to make adjacent on-the fly (OTF) strip maps of the Galactic Plane. The ultimate limiting factor for a larger pixel count array is the amount of available LO power. The results presented show that we have a scheme that can produce enough power to pump 4–10 HEB pixels for mapping of the Galactic Plane.

D.   NEW TECHNOLOGY

NTR number: 46567 — "A Quad-chip Double Balanced Frequency Tripler"

E.   FINANCIAL STATUS

The total funding for this task was $98,900, all of which has been expended.

F.   ACKNOWLEDGEMENTS

Help from Bertrand Thomas, Robert Lin, Seth Sin, Alex Peralta, and Alain Maestrini is acknowledged.

G.   PUBLICATIONS AND PRESENTATIONS

[A]   Choonsup Lee, J. Ward, R. Lin, E. Schlecht, G. Chattopadhyay, J. Gill, B. Thomas, A. Maestrini, I. Mehdi, and P. Siegel, "A Wafer-level Diamond Bonding Process to Improve Power Handling Capability of Submillimeter-Wave

Schottky Diode Multipliers," *Proceedings of the IEEE-IMS2009 Symposium*, Boston, Massachusetts, June 2009.

[B]   Imran Mehdi, John Ward, Alain Maestrini, Goutam Chattopadhyay, Erich Schlecht, Bertrand Thomas, Robert Lin, Choonsup Lee, and John Gill, "Broadband Sources in the 1–3 THz Range," *Proceedings of the 34th International Conference on Infrared, Millimeter, and Terahertz Waves*, Busan, Korea, September 2009.

H.   REFERENCES

None.
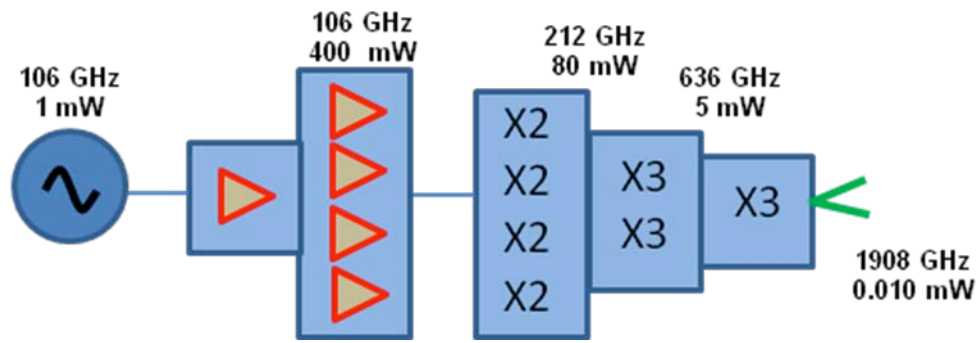
I.   FIGURES

Figure 1. This is the scheme that will be used to build the 1.9-THz source.
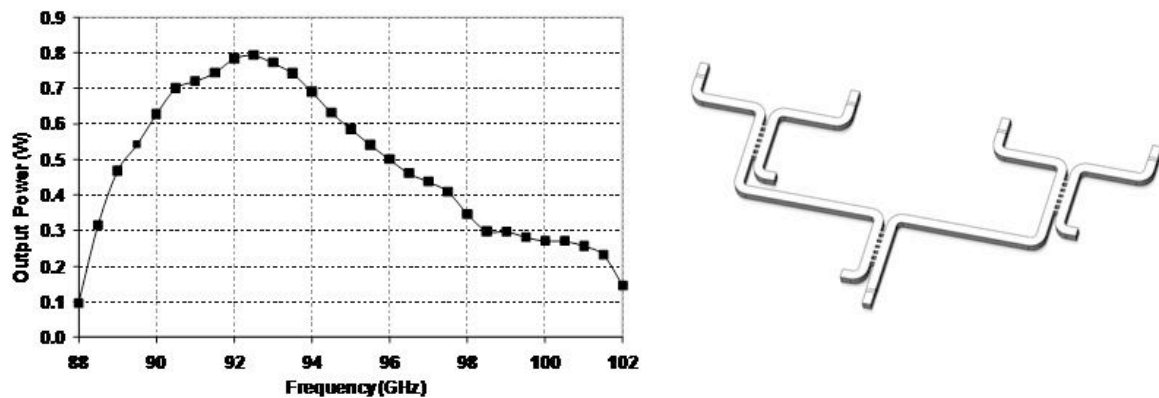
Figure 2. Schematic and measured results from the power-combined power amplifier module.
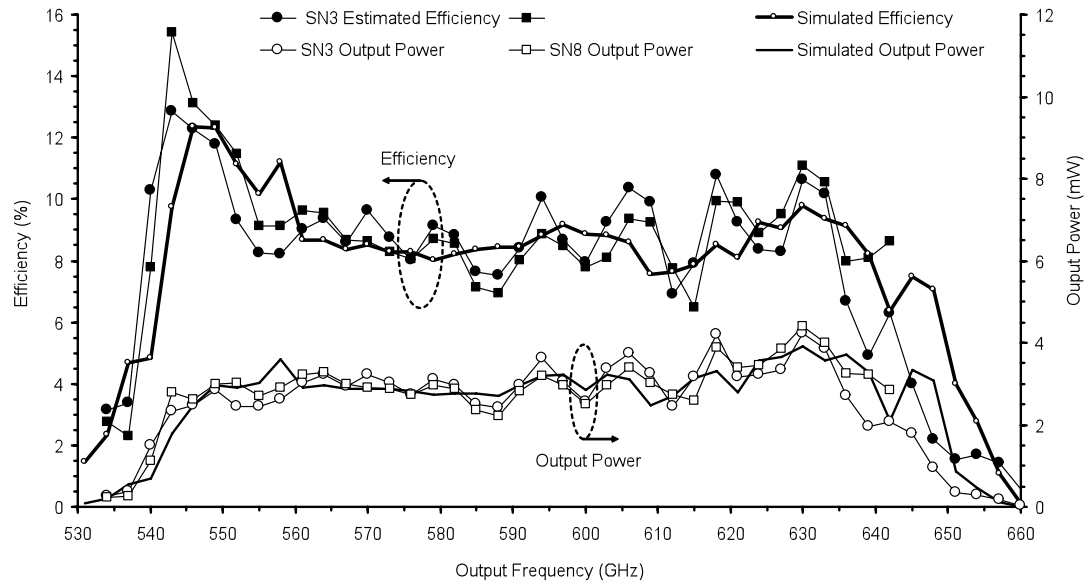
Figure 3. Measured data from a 600-GHz (×2×3) chain. Two of these multipliers will be power combined to enable >5-mW output power at 636 GHz.
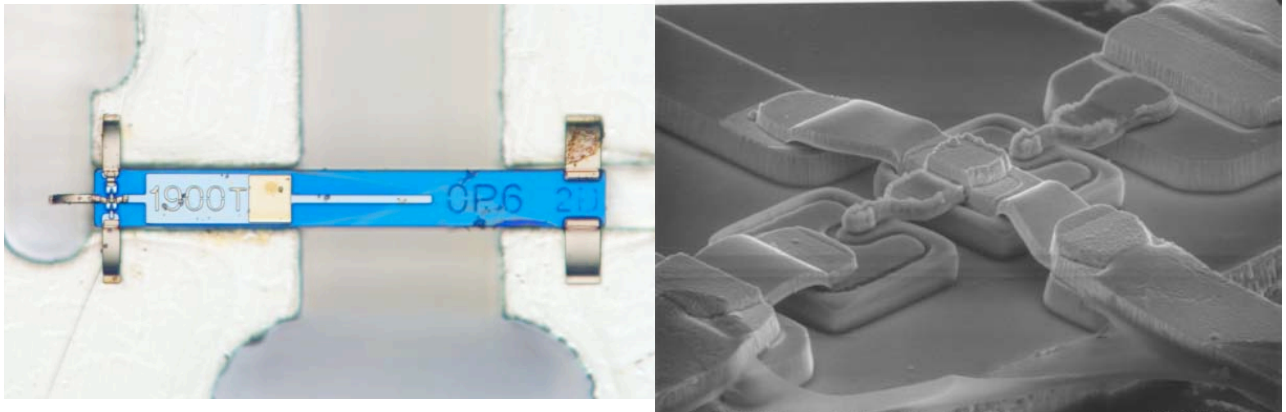
Figure 4. A 1900-GHz tripler chip with two anodes has been designed and fabricated (as shown in the left). A close-up of the anode area is shown on the right.

EARTH SYSTEM SCIENCE INFORMATION SYSTEM

THIN-FILM OPTICAL COATINGS FOR POLARIZATION SENSITIVE SYSTEMS

Nasrat A. Raouf (PI), Optical Section (383)
Russell Chipman (Co-PI), College of Optical Sciences, University of Arizona
Stephen McClain, College of Optical Sciences, University of Arizona
Paula Smith, College of Optical Sciences, University of Arizona

A.   OBJECTIVES

The objectives of this work were three. The first objective was to generate an algorithm to perform polarization ray tracing of a dielectric thin-film layer on a birefringent crystal substrate. This is the first step of possible algorithms to enable more complicated analysis of optical systems with multilayer-coated anisotropic materials for use in wave-plates, depolarizers, and optical filters. The second objective was to determine the appropriate thicknesses of a chosen dielectric layer (TiO_2) and then vacuum deposit this dielectric layer on the birefringent substrate. To show a large polarization effect for this demonstration, a calcite substrate was chosen to verify the algorithm developed for this task. The calcite crystal is a uniaxial anisotropic birefringent material with a known large difference between its n_o and n_e. The final objective was to compare the measured polarimetric performance of the TiO_2 dielectric thin film layer on the calcite substrate and the predicted performance from our modeling algorithm. To verify the algorithm of our model, the coated samples were measured on the Mueller Matrix Imaging Polarimeter (MMIP) facility of the University of Arizona.

B.   APPROACH AND RESULTS

The procedure for performing polarization ray tracing thorough uncoated birefringent media was established by Dr. S. McClain, University of Arizona, in his Ph.D. dissertation [1,2]. For the current work, the calculations of the reflectance and transmittance from the coating layer are treated through the continuity of fields in Maxwell's equations applied at the boundary between the thin-film layer and the substrate. In general, the current standard Abelias coating algorithm, cannot simply be applied twice, once for the ordinary and again for the extraordinary refractive indices. Due to interference effects between the thin-film layers, the amplitudes and phases of the individual waves and multiply-reflected partial waves must be carefully tracked and appropriately combined. Ray-splitting at the crystal interface further complicates the problem. The physics of this interaction have been treated by Mansuripur [2], but not in the context of polarization ray tracing. The task was to exploit existing physics and derive calculation methods that can be applied in a polarization ray tracing code for polarimetric optical systems.

The task of properly designing antireflective coatings or more specialized coatings for these materials is hampered by the lack of software that can handle a material that has multiple refractive indices with different dispersion curves.

This work addresses this shortcoming through design and testing of new algorithms. A dielectric layer was deposited on a birefringent substrate and performance characteristics were verified via laboratory measurements.

1.   Algorithm Development

We developed an algorithm model that allows us to trace rays through birefringent substrate materials with multilayer coatings. These coating layers can be any combination of isotropic, uniaxial crystal, or biaxial crystal materials. Coating characteristics are calculated from solving Maxwell's equations in each coating layer as well as considering the continuity of fields in Maxwell's equations applied at the boundary between the thin-film layers and at the substrate [3]. The algorithm model calculates the reflection and transmission from Jones matrices for the multilayer thin film stack. The diattenuation, retardance, and reflectance are calculated from the Jones matrix and compared to the measurement values achieved from the MMIP testing of the sample.

2.   Sample Fabrication

A set of ½-inch-diameter calcite substrates was procured commercially. These were coated with different layer thicknesses of TiO_2 and then tested in the MMIP setup for our demonstration.

To verify our algorithm model, we simulated the use of two design thicknesses of TiO_2 layer on a calcite substrate. This was to achieve in one case a high reflectance (293 nm of TiO_2 layer at 550 nm), and in the other case a low reflectance (352 nm of TiO_2 layer at 550nm).

We simulated the reflectance at normal incidence as the thickness of the thin film (TiO_2 layer) varies from a quarter-wave to three half-waves. The results showed a maximum at quarter-wave thicknesses and a minimum at half-wave thicknesses.

The MMIP setup was used to determine accurately the index of the TiO_2 at several wavelengths. This was accomplished by first depositing a 293-nm thickness of TiO_2 on a well-known index of refraction glass substrate. The measurements were compared to simulated diattenuation and retardance measurements. From these measurements, we were able to determine the index of refraction of TiO_2 at 450 nm to be 2.231, at 500 nm to be 2.147, and at 550 nm to be 2.139. These values differ from the expected values, which shifted the high-reflectance coating design of 293 nm to a thickness of five fourths lambda at 510 nm and the low-reflecting coating design of 352 nm to three halves lambda at 510 nm.

Three coating design thicknesses were used to coat the calcite substrate: a high-reflectance coating design with a 293-nm, a low-reflectance coating of 352-nm, as well as a 100-nm of TiO_2 thin film were then vacuum-deposited on the calcite substrates.

The diattenuation of each design was simulated over a range of wavelengths for an incident angle of 40º at different azimuthal angles (each plot represents a different azimuthal angle from 0º to 90º). These showed substantial diattenuation variation when compared to the uncoated calcite simulation and showed how each thin film layer influenced the diattenuation.

3.   Retardance and Diattenuation Measurements

The verification of the algorithm model was accomplished by comparing the measured data of the coated calcite in the MMIP setup with the simulated performance of the thin-film coatings using our algorithm model. The setup for the MMIP is described in [4].

The MMIP was set up in two configurations. In the first configuration, the incident beam is collimated to measure one pair of incident and azimuthal angles. In the second MMIP configuration, microscope objectives were added to the system to simultaneously measure a cone of incident angles. The second configuration results in an angle of incidence Mueller matrix image, where each subimage in the 4×4 grid of images represents a particular Mueller matrix element. Each pixel in a subimage represents a different incidence angle on the sample.

Each coated calcite sample was measured in reflection at three different azimuthal angle orientations (0º, 45º, and 90º) with an incident angle varying from 0º to 70º in increments of 5º. These measurements were performed in collimated light at 450-nm, 500-nm, and at 550-nm wavelengths. Each sample was also measured with the microscope objective configuration to obtain a Mueller matrix for a cone of incident angles. The thickness values were optimized for each sample using the optimized indices from the witness sample and the measurement results.

The resulting optimized thicknesses were 114 nm, 298 nm, and 349 nm. Figure 1 compares the diattenuation and retardance at three different azimuthal angle orientations (0º, 45º, and 90º) with the collimated light measurements and simulation (using the optimized indices and thicknesses) for the 349-nm sample at 500-nm wavelength. The measured data points show good agreement to the simulation. Similarly good agreement between simulation predications and measurement results was demonstrated for the other samples and wavelengths, but page limitations prevent us from showing additional data.

Microscope objectives were used in the polarimeter produce angle of incidence (AOI) Mueller matrix image data, where each pixel represents a different set of incident and azimuthal angles.

Each calcite sample was measured in reflection at 500nm with a central incident angle of 45º. Figure 2 (top) shows the measured angle of incidence Mueller matrix image data for the calcite sample with a 298-nm layer of TiO_2 with index of 2.147. After optimizing the refractive index and the thickness of the TiO_2 thin-film layer, a more accurate index at 2.176 and a thickness of 297 nm were used to simulate the image data shown on the bottom of Figure 2. Very good agreement is seen between the simulation and measurements.

4.   Further Analysis and Modeling

We have extended the model to include a simulation of an antireflection (AR) coating design and its performance on a calcite substrate. Using the biaxial option of the algorithm model, a two-layer antireflection coating was designed for a calcite substrate at 550 nm. The goal of this design was to minimize the average reflectance for an azimuthal angle at 45º. The AR coating design started with two quarter-wave-thick layers and was optimized for incident angles from −10º to 10º. The optimum indices for the two layers were determined to be 2.135 (for the layer closest to the substrate) and 1.845, and the optimized thicknesses were quarter wave or 59.4 nm (for the layer closest to the substrate) and 74.5 nm. This design results in a reflectance of 0.078% at normal incidence and 0.084% at ±10º. The reflectance resulting from the two-layer AR coating design for incident angles was calculated from −60º to 60º and shows improvement from the results of a single-layer AR coating design where the optimized layer had an index of 1.5 and a quarter-wave-layer thickness of 91.7 nm. The reflectance resulting from the single layer had 3.2% reflectance at normal incidence and 3.2% reflectance at ±10º. This AR coating design shows the biaxial algorithm model can be applied to designing coatings by optimizing the merit functions for a birefringent substrate.

C.   SIGNIFICANCE OF RESULTS

In this research, we have derived and validated optical analysis algorithms for predicting the performance of thin-film coatings on birefringent substrates. Since commercially available optical design software does not provide this capability, it has not been possible to validate the behavior of optical designs containing coated birefringent elements, and it has not been possible to optimize the performance of birefringent elements in polarization critical optical systems. This capability can provide JPL with a competitive advantage for designing sensors, cameras, and radiometers for which polarization performance is a driving performance requirement, including but not limited to MSPI cameras.

D.   FINANCIAL STATUS

The total funding for this task was $100,000, of which $77,259 has been expended.

E.   ACKNOWLEDGEMENTS

F.   PUBLICATIONS

None.

G. REFERENCES

[1] Stephen C. McClain, Lloyd W. Hillman, and Russell A. Chipman, "Polarization Ray Tracing in Anisotropic Optically Active Media. I. Algorithms," *J. Opt. Soc. Am. A* **10** (1993): pp. 2371–2382.

[2] Stephen C. McClain, Lloyd W. Hillman, and Russell A. Chipman, "Polarization Ray Tracing in Anisotropic Optically Active Media. II. Theory and Physics," *J. Opt. Soc. Am. A* **10** (1993): pp. 2383–2393.

[3] M. Mansuripur, "Analysis of Multilayer Thin-Film Structures Containing Magneto-Optic and Anisotropic Media at Oblique Incidence Using 2×2 Matrices," *Journal of Applied Physics* **67** (May 15, 1990): pp. 6466–6475.

[4] Paula K. Smith and Russell A. Chipman, "Characterizing Dielectric Tensors from Angle-of-Incidence Mueller Matrix Images," *Proc. SPIE* 6682, 66820N, 2007.

H. FIGURES
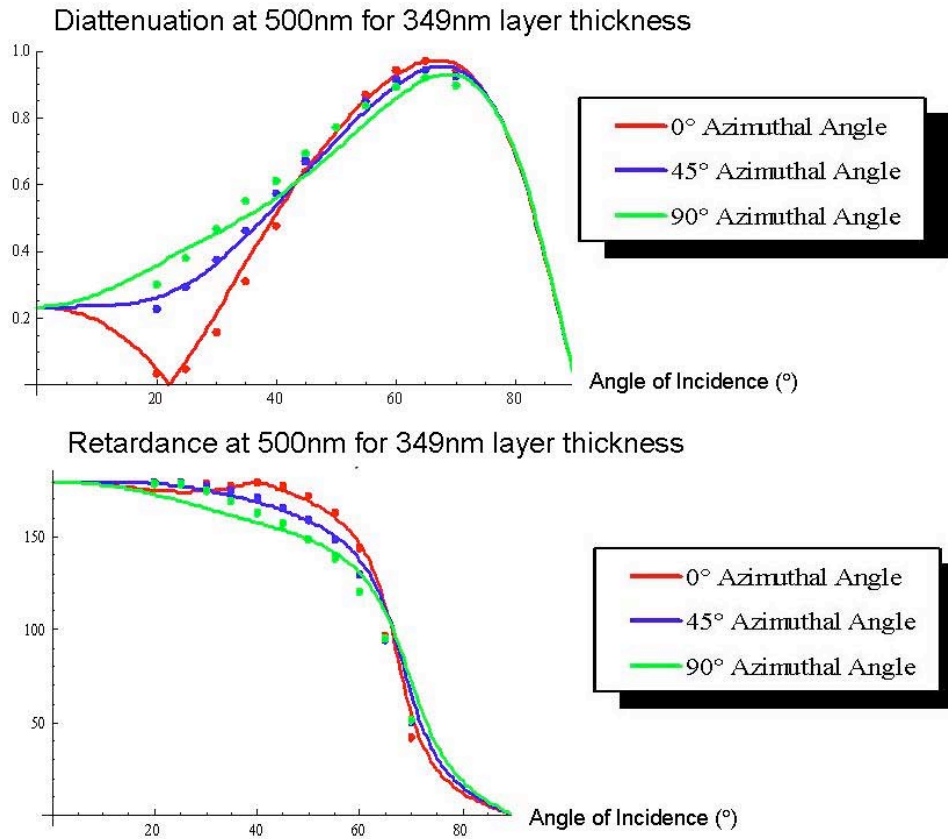
Figure 1. Compare the diattenuation for sample with 349-nm layer of TiO_2 at 500 nm for three different azimuthal angle orientations (0º, 45º, and 90º). The collimated light measurements are shown as dots and simulation is shown as lines.
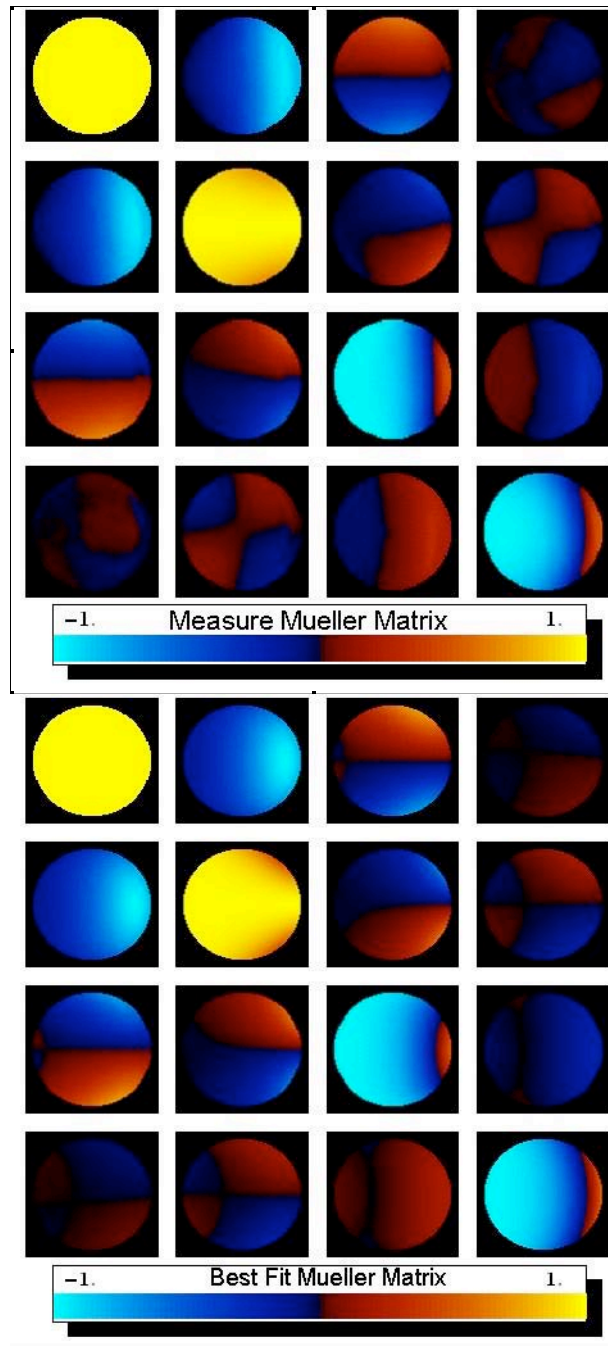
Figure 2. Measured (top) and simulated (bottom) angle of incidence Mueller matrix for the calcite sample with 298-nm layer of TiO_2 measured in reflection at 500 nm with M_{00} normalized to 1. The cone of measured angles is ±30º with the center pixel corresponding to an incident angle of 45º.

FROM DETECTION TO IMAGING: ENHANCING THE SCIENCE RETURN OF DIRECT IMAGING EXOPLANET MISSIONS VIA PRIOR ASTROMETRIC PLANET DETECTION

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1335

Joseph Catanzarite (PI), Optics Section (383)
Michael Shao, Interferometry Center of Excellence (302)
Renu Malhotra, Department of Planetary Sciences, University of Arizona

A.   OBJECTIVES

This work was motivated by two objectives: (1) quantify the improvement in the science return of a mission to detect exo-Earths by direct imaging that can be achieved with prior astrometric detection and orbit characterization; and (2) develop algorithms to help reject false planet detection by identifying dynamical instability.

B.   APPROACH AND RESULTS

1.   The Synergy of Direct Imaging and Astrometry for Detection of exo-Earths
         Personnel: PI Joseph Catanzarite; Co-I Michael Shao

In this project, we simulated the planet-finding capability of two external occulter mission concepts: (1) the Occulting Ozone Observatory, and (2) the James Webb Space Telescope plus a starshade. We developed a coronagraph mission modeling code (MATLAB) that included a realistic model of planet mass and semimajor axis distributions that is consistent with current observational data.

We showed that a single image of an exo-Earth can improve the positional precision of the orbital solution obtained by a prior astrometric mission by a factor of 10.

We showed that in the absence of prior astrometric data, a single image of an exo-Earth candidate is 50% likely to be a false alarm, i.e., a planet that is not an exo-Earth. In order to mitigate false alarms, it is necessary to image the planet at least four times in order to fit and verify the orbit. In practice, each planet must be observed more than four times in order to get four images, since the planet is not always visible. Even if the orbit can be fitted, the mass is undetermined and an exo-Earth candidate may turn out to be a Neptune — a false alarm.

We found that a prior astrometric mission (with submicroarcsecond precision) would enhance the exo-Earth yield of the external occulter missions by a factor of about 4 to 5, if the occurrence frequency of exo-Earths is between 0.1 and 0.3. If the frequency of exo-Earths is

0.1, then without a prior astrometry mission, a JWST plus starshade mission has 40% chance of finding no exo-Earths.

2.   Dynamical Stability Study
Personnel: Co-I Renu Malhotra (Lunar and Planetary Laboratory, University of Arizona) was assisted by UA graduate student Youngmin JeongAhn in this effort. The numerical calculations were carried out on a computer cluster in Malhotra's research laboratory.

In the present project, we identified analytical and numerical tools to test for the dynamical stability of the model-fits, and we used these to investigate a subset of the synthetic systems that were identified as problematic cases in the prior double-blind study. We used these investigations to gain insight into how such data and analysis might yield improved models.

Analysis of high-precision position and velocity data of stars potentially allows the detection of planetary systems harbored by those stars. A basic step in the analysis is the fitting of Keplerian orbital models to extract the parameters of the companions. The double-blind study done previously by the PI and collaborators developed and tested state-of-the-art tools for such model-fitting. One of the challenges that emerged from that work was how to address cases of non-unique solutions with similar goodness-of-fit measures. It was noted that short-term (a few-to-tens of orbital periods) and long-term orbital stability are expected to be characteristics of real multiple-planet systems, but this property is not included in the Keplerian model fitting tools which do not take account of the mutual gravitational interactions of the planetary companions. In the present project, we identified analytical and numerical tools to test for the dynamical stability of the model-fits, and we used these to investigate a subset of the synthetic systems that were identified as problematic cases in the prior double-blind study. We used these investigations to gain insight into how such data and analysis might yield improved models.

a.   Analytical tools for dynamical stability of multiple systems

1)   Orbit crossing test: A system fails this test if the periastron-to-apoastron distance range of a planet overlaps the periastron-to-apoastron distance range of a neighboring planet. A system that fails this test is dynamically unstable on orbital timescales, unless a mean motion or secular resonance condition exists to force a mutual stabilizing phase relationship between the two planets.
2)   Resonance overlap test: A system fails this test if the semimajor axis of a planet falls within the chaotic zone of a neighboring planet. (The chaotic zone arises from the overlapping of neighboring first-order mean motion resonances [1,2]). The chaotic zone of a planet on a circular orbit is an annular region centered on the planet's orbit, of half-width $\Delta a = 1.5\,(m_p/m_s)^{2/7}a$, where a is the planet's semimajor axis, and m_p and m_s are the mass of the planet and star, respectively. For eccentric planetary orbits, a formal estimate of the resonance overlap zone is not available; informally, we may extrapolate the above formula by defining the inner and outer edges of the chaotic zone as follows: $a_{inner} = q(1-1.5\,(m_p/m_s)^{2/7})$, $a_{outer} = Q(1+1.5\,(m_p/m_s)^{2/7})$, where q and Q are the planet's periastron and apoastron distances, respectively.
3)   Note that this is an approximate criterion. (This owes to the non-integrability of the three-body problem.) A system that fails this test is dynamically unstable on

orbital timescales, but a system that fails this test "marginally" may be quasi-stable for fairly long periods of time, exhibiting only low-amplitude chaotic evolution. Herein lies much of the challenge of defining a practical criterion for stability of planetary systems.

b.   Numerical tools for dynamical stability of multiple systems

Perhaps the definitive test for dynamical stability is by means of numerical integration, but it bears a high cost in computational time and effort. In the simplest model, we need to solve a set of coupled ordinary differential equations. Historically, the challenge with planetary systems has been the long computational times required for even moderate length of time propagation of the solution with only moderate numerical accuracy. Recent advances in numerical methods for symplectic Hamiltonian systems have been applied very successfully to solar system dynamics, and allow for much faster orbit integrations. Even so, examining a large parameter space of model-fits of data remains a very challenging computational problem. We implemented the second-order mixed-variable symplectic integrator of [3], adopting algorithms described in [4]; our code is written in FORTRAN-77 and is called MVS2. We used this to study the dynamical stability on timescales up to 10 myr of the model systems provided by PI Catanzarite. The timescale 10 myr was chosen as a compromise: strong instabilities are expected to manifest on a few to a few hundreds of orbital periods; weak instabilities manifest on gigayear timescales but are not relevant as a criterion for model-fitting; intermediate timescales of 108 yr are relevant to identifying non-viable models; however, the computational load for examining more than a few cases becomes prohibitive.

c.   Results

•   A total of 27 different model systems were studied. Two versions of each case were examined: the "original" system and the "fitted" system (fits based on simulated observations of those systems, all data supplied by Joseph Catanzarite).

•   Somewhat surprisingly, we found that eight of the "original" systems failed one or more of the dynamical stability tests. This is surprising because it was previously assumed that these systems were stable by construction.

•   Of the "fitted" systems, we found 12 systems that failed the analytical dynamical stability tests; all but one of these was confirmed unstable by numerical analysis. The one exception avoids instability due to an apsidal resonance condition.

•   Careful comparison of the "fitted" systems that failed the dynamical stability tests and their corresponding "original" systems shows that often the model fits have fewer planets and significantly larger orbital eccentricities. The reason for the fewer planets in the models vs. original is most often the very low masses of the non-detected planets. We spent much effort to understand the reason for the often spuriously large orbital eccentricities, but this remains unclear. This problem has been recognized in the literature for some time now; it is often attributed to undetected long-period planets in the system, but our examination of the present data set does not confirm this explanation.

•　　　It is of note that only three of the eight unstable "original" systems are also found unstable in its "fitted" model. In the other cases, the "fitted" models are (spuriously) stable for a variety of reasons: the number of detected planets is smaller, or (in one case) an apsidal resonance condition in the fitted model assures stability. The latter case is quite remarkable, as it serves as a warning that model fits may lead us astray in inferring exotic dynamical evolution where none may actually exist. (We note that dynamical theories link an apsidal resonance condition to a history of planet–planet scattering or planet migration, e.g., [5,6]).

d.　　　Future directions

•　　　There is an urgent need to derive an analytical "resonance overlap" criterion for eccentric planetary orbits (or to prove the validity of our informal extension of this criterion from circular orbits). Furthermore, the "resonance overlap" criterion is valid in the limit of very small ratio of planet-to-stellar mass; exoplanet systems with planet masses near the brown dwarf mass range challenge us to examine the validity of this criterion at this range of planet-to-mass ratios.

•　　　The fast symplectic integrator for planetary system dynamics needs to be extended to include the effects of general relativity. This is of relevance in cases of massive, short-period planets, where general relativistic precession can be important for dynamical stability in multiple-planet systems.

C.　SIGNIFICANCE OF RESULTS

1.　　　The Synergy of Direct Imaging and Astrometry for Detection of exo-Earths

Our stated objective for this part is "to quantify the gain in science return from a direct imaging mission, when the masses and orbits of potentially habitable planets have been previously estimated from astrometric measurements."

We have done this, and showed that an astrometric precursor mission would dramatically enhance the exo-Earth yield of two leading direct imaging mission concepts.

Our work is the first to model imaging surveys with realistic planet semimajor axis and mass distributions for the target stars. In particular, we investigated the effects of confusion due to multiple planets on the science yield.

Impact on future JPL missions and programs: Our work highlights the fact that existing external occulter and internal coronagraph concepts are very inefficient at finding exo-Earths when η_{Earth} (the occurrence rate of exo-Earths) is in the expected range of 10% to 30%. In this case, a realistic program to find and take the spectra of exo-Earths orbiting nearby stars should be preceded or accompanied by an astrometric mission to identify the exo-Earths and measure their orbits and masses.

In addition, the Appendix of publication [A] contains an assessment of the requirements of speckle subtraction in order to detect exo-Earths with a future coronagraph mission.

2.   Dynamical Stability Study

Our stated objective for this part is to "Develop algorithms to reduce false planet detections when solving planetary orbits of multiple-planet systems from astrometric and RV data."

We are less optimistic now about using dynamical stability as a criterion to improve model fits than when we began this project. The value of demanding dynamical stability of a fitted system lies in flagging problematic cases, and to motivate and justify efforts for more observations of such systems. However, the parameter space for exploring "nearby" stable models is prohibitively large, as it must include not only the uncertainties in the parameters of detected planets but also a very poorly restricted number of additional planets' parameters. There is some suggestion in the exoplanets literature that planetary system configurations have a propensity to be "near the edge of dynamical stability." This is an interesting suggestion and deserves further theoretical and modeling studies, but we think that it is as yet too undeveloped and unproven to be of practical use in data analysis and model-fitting of radial velocity/astrometric observations of exoplanet systems.

D.   FINANCIAL STATUS

E.   ACKNOWLEDGEMENTS

F.   PUBLICATIONS AND PRESENTATIONS

[A]   Michael Shao, Joseph Catanzarite, and Xiaopei Pan, "The Synergy of Direct Imaging and Astrometry for Orbit Determination of exo-Earths," *Astrophysical Journal* **720** (September 1, 2010): pp. 357–367.

[B]   Joseph Catanzarite and Michael Shao, "Viability of a Mission to Characterize Exo-Earths Using JWST plus a Starshade External Occulter," *Proceedings of the Astronomical Society of the Pacific*, submitted, under referee review.

[C]   Joseph Catanzarite, "Measuring the Orbits of Exo-Earths in Multiple-Planet Systems: The Synergy of Direct Imaging and Astrometry for Detection of exo-Earths," talk presented at the American Astronomical Society, AAS Meeting #216, May 23–27 2010, Miami Florida, #311.04.

[D]   Joseph Catanzarite, "The Synergy of Direct Imaging and Astrometry for Detection of exo-Earths," talk to be presented at In the Spirit of Lyot 2010 Paris, Direct Detection of Exoplanets and Circumstellar Disks, October 25th–29th, 2010.

G.   REFERENCES

[1]   B. V. Chirikov, "A Universal Instability of Many-Dimensional Oscillator Systems," *Phys. Rep.* **52** (1979): pp. 263–379.

[2]   J. Wisdom, "The Origin of the Kirkwood Gaps — A Mapping for Asteroidal Motion Near the 3/1 Commensurability," *AJ* **87** (1982): p. 577.

[3]   J. Wisdom and M. Holman, "Symplectic Maps for the N-Body Problem," *AJ* **102** (1991): p. 1528.

[4]   P. Saha and S. Tremaine, "Symplectic Integrators for Solar System Dynamics," *AJ* **104** (1992): p. 1633.

[5]   E. Chiang and N. Murray, "Eccentricity, Excitation, and Apsidal Resonance Capture in the Planetary System Upsilon Andromedae," *ApJ* **576** (2002): pp. 473–477.

[6]   R. Malhotra, "A Dynamical Mechanism for Establishing Apsidal Resonance," *ApJ* **575** (2002): pp. L33–L36.

EDUCATION

**DEVELOPMENT OF A GLOBAL WATER CYCLE MODELING
AND ASSIMILATION SYSTEM**

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1311

Eni G. Njoku (PI), Climate Oceans and Solid Earth Science Section (324)
Eric F. Wood (Co-PI), Department of Civil and Environmental Engineering,
Princeton University

A. OBJECTIVES

The objective of the task was the development of initial elements for a global water cycle modeling and assimilation system testbed that can provide a "seamless suite of water cycle information." The task aimed at a system that would merge hydrologic data and model predictions across sensors, models, and scales that could then be used for water cycle prediction and in decision support models.

The above objective addresses a longer-term NASA science and applications research question: *Can global remote sensing products, from current and anticipated space sensors, provide the basis for scientific water cycle studies and water resources applications (e.g., flood and drought monitoring), particularly in those parts of the world where in situ networks are too sparse to support more traditional methods of hydrologic prediction?*

B. APPROACH AND RESULTS

1. General approach

The general approach consisted of (i) <u>develop the system framework</u> for the merging of data and predictions across sensors and models, (ii) <u>construct a prototype testbed</u> for the system, and (iii) <u>demonstrate that the testbed</u> can be used to assess current and proposed (i.e., future "decadal survey" and NPOESS) satellite sensors in providing information needed for water cycle studies and water resources applications. Building the first project was organized as a number of tasks that developed around the recognition that the proposed system will integrate the strengths at Princeton in global water cycle modeling, hydrologic remote sensing, and data assimilation with the strengths at JPL in remote sensing retrievals, sensor design and analysis, and orbit/sensor sampling and simulation. The project offers the unique opportunity to bring together these existing capabilities into an "integrated Earth System Science information system."

2.   Results from the project

a.   An overview of the developed framework for the system is provided in Figure 1. The left side of the figure represents the "water cycle information" or monitoring functions that would take remote sensing observations (say radiation, precipitation, etc.) that can be used as input variables to a high-resolution land surface "Earth System" model as well as remote sensing observations like soil moisture, river stage or vegetation state that can be assimilated into the land surface model to provide merged information appropriate for decision making.

The right side of Figure 1 is a classical observing system simulation experiment (OSSE) that utilizes many elements from the water information side — particularly the land surface model that can provide the water cycle state that is then coupled to forward models to provide top-of-atmosphere (TOA) observations, which are then sampled depending on the characteristics of the sensors and antenna, as well as the sampling characteristics of the satellite orbit. The simulated satellite measurements are then used to simulate retrieved water cycle variables that are used to evaluate the satellite and sensor characteristics for retrieval accuracy, their usefulness in decision-making, and so forth.

b.   The development of the detailed land surface modeling is based on the variable infiltration capacity (VIC) land surface model (LSM). As discussed above, it is used in the information system to merge model predictions with the satellite-based retrievals in a data assimilation system that provides predictions of water cycle variables. It is also used in the OSSE system to provide water cycle predictions that would be inputs into the forward models of the OSSE system. LSMs represent the physical Earth System processes that control the land surface water balance. In off-line simulations, LSMs are typically forced by precipitation, radiation, and surface meteorological variables. The VIC LSM of [1] is one of the LSMs in the NOAA/NASA North American Land Data Assimilation System (NLDAS) [2] and is being used in the Surface Water and Ocean Topography (SWOT) virtual mission OSSE. Compared to other LSMs, VIC's distinguishing features are its representation of subgrid variability in soil moisture storage capacity, to which surface runoff is related, its parameterization of base flow from a lower soil moisture zone as a nonlinear recession, and orographic precipitation and temperature lapse rates resulting in more realistic mountain hydrology. River routing is represented with a linear transfer function approach and has been recently updated to represent water management effects [3], including reservoirs and irrigation, features important for SWOT and SMAP missions.

The VIC model was run at 5-km spatial resolution over Africa for 2003. The underlying physiographic parameters required by the model are topography, soils, and vegetation. Topography is taken from the USGS Hydro1K geographic database, which is derived from the GTOPO30 DEM. Soils and vegetation characteristics and their spatial distribution were specified following [4]. Soil textural information and bulk densities were derived by combining the 5-min Food and Agricultural Organization–United Nations Educational, Scientific, and Cultural Organization (FAO–UNESCO) digital soil map of the world [5] with the World Inventory of Soil Emission Potentials (WISE) pedon database [6]. The remaining soil characteristics, such as porosity, saturated hydraulic conductivity, and the exponent for the unsaturated hydraulic conductivity equation were based on [7]. Vegetation types were taken

from the Advanced Very High Resolution Radiometer (AVHRR)–based, 1-km, global land classification of [8]. Vegetation parameters such as height, and minimum stomatal resistance, were assigned to each vegetation class based on a variety of sources described in [4]. Monthly values of leaf area index were based on [9] and were kept constant from year to year.

Meteorological forcing data are based on the global dataset of [10], which has been updated to 2008. The temperature, humidity, pressure, and longwave radiation data were downscaled in space by lapsing with elevation. To obtain realistic precipitation spatial patterns, the original coarse precipitation was replaced by remote sensing based estimates from the Global Satellite Mapping of Precipitation (GMSMAP; [11]), which is based on a multisatellite blend of microwave and infrared retrievals to give 0.1-degree hourly data, that are also available in near real time.

c.   Results were achieved for testing the OSSE component of the system (the right side of Figure 1) using the 5-km VIC LSM simulations described above. The period simulated was 2003 over Africa. The OSSE experiments focus on the radiative transfer process where land surface states, e.g., soil moisture (SM) and surface temperature (Ts), determine the electromagnetic signals (Tb) detectable by spaceborne sensors, as well as the reverse process of retrieving hydrologic states from satellite measurements. The set of "true" land surface states were generated by the VIC LSM (as described above) and the states passed to a forward radiative transfer model (RTM) to obtain the "true" Tb. The measurement of these brightness temperatures by the satellite sensor/antenna system (using specific design criteria) is simulated with added measurement errors to the "true" Tb. For this task, the forward RTM is the Land Surface Microwave Emission Model (LSMEM). LSMEM calculates the brightness temperature (Tb) at the top of atmosphere given the conditions of soil (soil moisture SM, surface temperature Ts, soil texture, etc.), vegetation (thickness, structure, temperature, etc.), and atmosphere (temperature, pressure, humidity, etc.) as described in [12] and [13]. Then the RTM is run in inverse mode to retrieve land surface states, allowing an assessment of sensor/satellite design (e.g., footprint size, scanning pattern, overpass times, and frequency), instrument errors, parameter errors, etc., on retrieval accuracy.

The OSSE here were conducted following these steps: (1) Run VIC at 5 km (0.05°) to obtain a set of synthetically true SM and Ts; (2) use these within LSMEM in its forward mode to obtain synthetically true Tb; (3) aggregate Tb to 15 km (0.15°) and 45 km (0.45°) to match the size of potential antenna footprints and add random sensor errors; (4) sample to represent the satellite orbits (swaths) in space and time; (5) retrieve SM using LSMEM in its retrieval mode; and (6) compare to the retrievals to the "true" SM and calculate errors and assess their values in land surface data assimilation. Figure 3 shows an example result.

C.   SIGNIFICANCE OF RESULTS

The core elements of the proposed system have been developed and tested, meeting the task objective to develop the initial elements for a global water cycle modeling and assimilation system testbed. This core can be expanded to include satellite observations for other water cycle variables (precipitation, water vapor, snow, surface water, evapotranspiration) or vegetation states to offer monitoring — the "seamless suite of water cycle information" — or as a system

for assessing new satellite sensors as new forward models become available. It is envisioned that this system can be built on and expanded at JPL to assess potential water cycle satellite sensors or to evaluate how current sensors can be used for water cycle monitoring in a virtual mission setting.

D. FINANCIAL STATUS

The total funding for this task was $99,910, of which $84,800 has been expended.

E. PUBLICATIONS

None.

F. REFERENCES

[1] X. Liang, D. P. Lettenmaier, E. F. Wood, and S. J. Burges, "A Simple Hydrologically Based Model of Land Surface Water and Energy Fluxes for General Circulation Models," *J. Geophys. Res*. **99** (1994): pp. 14415–14428.

[2] K. E. Mitchell, et al., "The Multi-Institution North American Land Data Assimilation System (NLDAS): Utilizing Multiple GCIP Products and Partners in a Continental Distributed Hydrological Modeling System," *J. Geophys. Res*. **109** (2004). 10.1029/2003JD003823

[3] I. Haddeland, T. Skaugen, and D. P. Lettenmaier, "Anthropogenic Impacts on Continental Surface Water Fluxes," *Geophys. Res. Lett*. **33** (2006): L08406. doi:10.1029/2006GL026047

[4] B. Nijssen, G. M. O'Donnell, D. P. Lettenmaier, D. Lohmann, and E. F. Wood, "Predicting the Discharge of Global Rivers," *J. Climate* **14** (2001): pp. 3307–3323.

[5] World Resources Institute, "Soils Classification for North and South America (FAO)," http://www.wri.org/publication/content/7922.

[6] N. H. Batjes, "A Homogenized Soil Data File for Global Environmental Research: A Subset of FAO, ISRIC, and NCRS Profiles," *Rep. 95/10,* International Soil Reference and Information Centre, Wageningen, Netherlands, 50 pp., 1995 [available from ISRIC, P.O. Box 535, 6700 AJ Wageningen, Netherlands].

[7] B. J. Cosby, G. M. Hornberger, R. B. Clapp, and T. R. Ginn, "A Statistical Exploration of the Relationships of Soil Moisture Characteristics to the Physical Properties of Soils," *Water Resour. Res*. **20** (1984): pp. 682–690.

[8] M. C. Hansen, R. S. DeFries, J. R. G. Townshend, and R. Sohlberg, "Global Land Cover Classification at 1 km Spatial Resolution Using a Classification Tree Approach," *Int. J. Remote Sens.* **21** (2000): pp. 1331–1364.

[9] R. B. Myneni, R. R. Nemani, and S. W. Running, "Estimation of Global Leaf Area Index and Absorbed PAR Using Radiative Transfer Models," *IEEE Trans. Geosci. Remote Sens.* **35** (1997): pp. 1380–1393.

[10] J. Sheffield, G. Goteti, and E. F. Wood, "Development of a 50-yr, High-Resolution Global Dataset of Meteorological Forcings for Land Surface Modeling," J. Climate **13** (2006): pp. 3088–3111.

[11] K. Aonashi, J. Awaka, M. Hirose, T. Kozu, T. Kubota, G. Liu, S. Shige, S. Kida, S. Seto, N. Takahashi, and Y. N. Takayabu, "GSMaP Passive, Precipitation Retrieval Algorithm: Algorithm Description and Validation," *J. Meteor. Soc. Japan* **87A** (2009): pp. 119–136.

[12] M. Drusch, E. F. Wood, and T. J. Jackson, "On the Impact of the Atmosphere in Passive Microwave Remote Sensing of Land Surface Parameters," in *Remote Sensing and Hydrology 2000*, M. Owe, K. Brubaker, J. Ritchie, and A. Rango (editors), IAHS Publ. 267, IAHS Press, Wallingford, Oxon, UK, 2001: pp. 400–404.

[13] Huilin Gao, Eric F. Wood, Matthias Drusch, Wade Crow, and Thomas J. Jackson, "Using a Microwave Emission Model to Estimate Soil Moisture from ESTAR Observations During SGP99," *J. Hydrometeor* **5** (2004): pp. 49–63.

G.   FIGURES
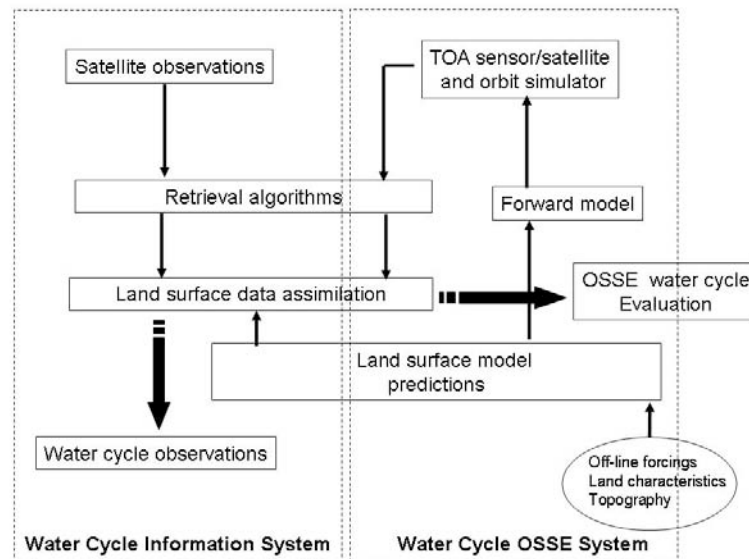
Figure 1. Element of the water cycle modeling and assimilation system.
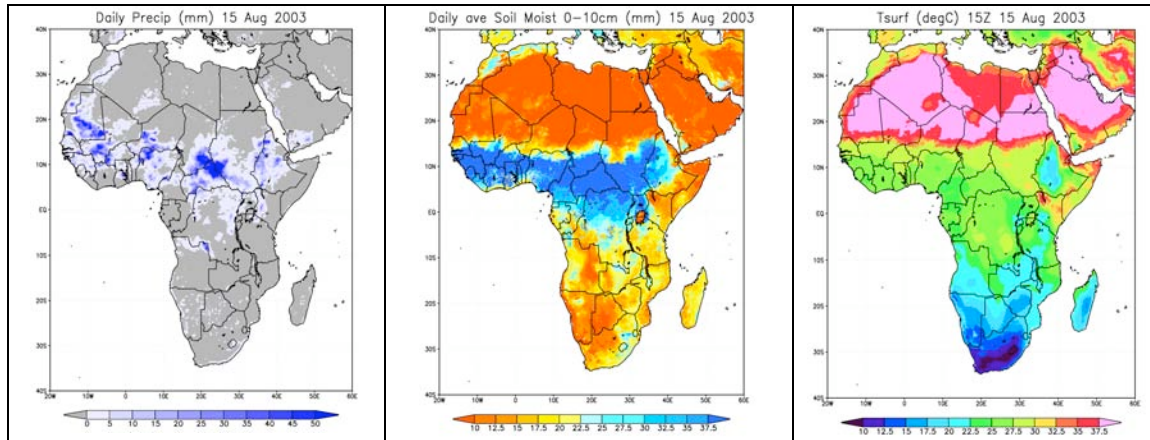
Figure 2. Example maps of daily mean precipitation forcing to the VIC model and VIC predicted upper layer soil moisture and surface temperature for 15 August 2003.
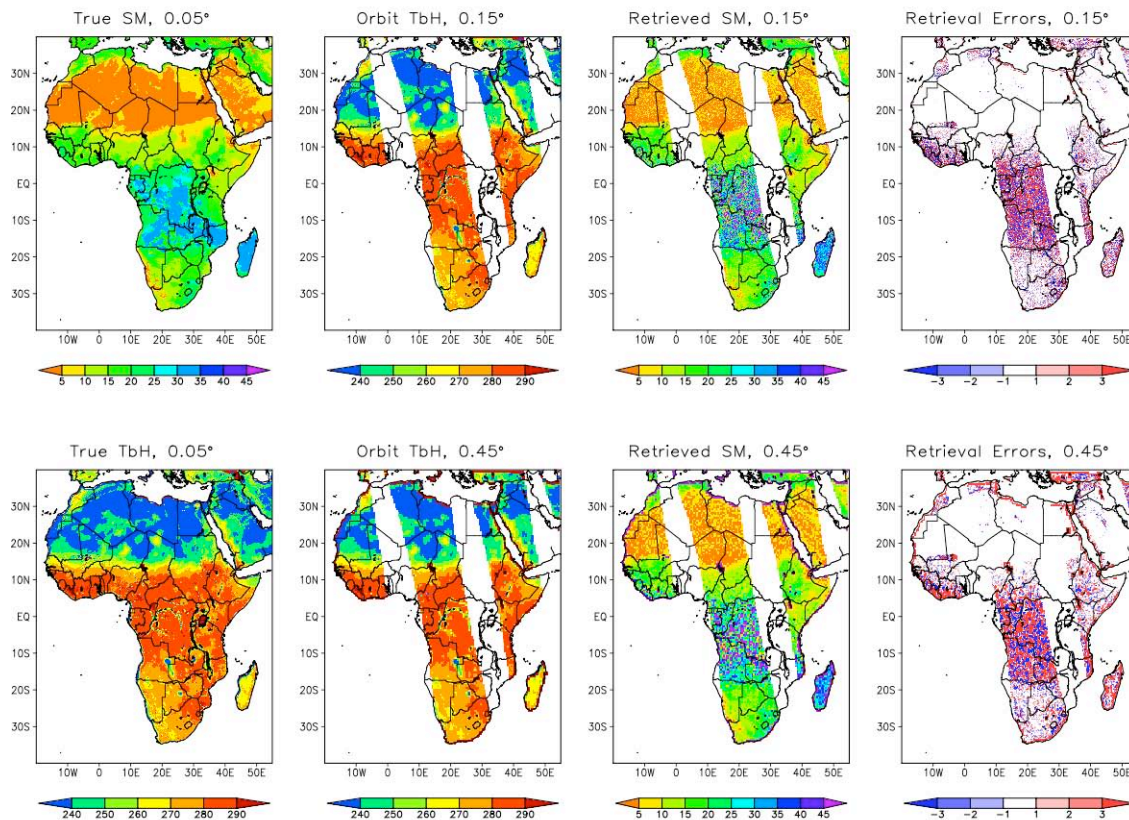
Figure 3. OSSE experiment results for 31 January 2003.

INTEGRATING WRF SIMULATIONS AND REMOTE SENSING OBSERVATIONS IN THE JPL HURRICANE FORECASTING AND ANALYSIS SYSTEM

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1316

Svetla M. Hristova-Veleva (PI), Radar Science and Engineering Section (334)
Hui Su, Microwave Atmospheric Science Section (328)
Robert G. Fovell, Department of Atmospheric and Oceanic Sciences and Joint Institute for Regional Earth System Science and Engineering, University of California at Los Angeles
Yu Gu, Department of Atmospheric and Oceanic Sciences and Joint Institute for Regional Earth System Science and Engineering, University of California at Los Angeles

A.   OBJECTIVES

Today there are still many unanswered questions about the physical processes that determine hurricane genesis, intensity, track and impact on the large-scale environment. Improving the forecasting of hurricane genesis and evolution — particularly sudden intensification and weakening — remains a challenge for the operational and research communities, and a significant amount of work remains to be done in validating hurricane forecast models, understanding their sensitivities, and improving their parameterizations. None of this can be accomplished without the fusion of a comprehensive set of multiparameter observations and model simulations relevant to both large-scale and storm-scale processes in the atmosphere and in the ocean.

The JPL Tropical Cyclone Information System (TCIS) provides a unique prototype to connect the state-of-the-art satellite and in situ datasets with model simulations for hurricane studies. The TCIS architecture was first designed with funding from the JPL's Innovative Spontaneous Concepts (ISC) program in 2007. The TCIS development is now funded as part of NASA's Hurricane Science Research Program.

The TCIS provides a wealth of multi-instrument multisatellite and in situ atmospheric and oceanic datasets, with both large-scale and storm-scale measurements of the dynamical, thermodynamic, and microphysical properties of hurricanes (Figure 1, left panel). We have also begun the development of an analysis system that will allow on-line evaluation of the statistical properties of the observed systems (Figure 1, right panel). Compared to other existing hurricane data systems, one of the unique features of the JPL TCIS is its close connection with state-of-the-art hurricane model simulations, so that the remote sensing data can be directly applied to evaluate model simulations, guide model improvements, and ultimately lead to improved hurricane forecasts. In this study, we focus on developing strategies for model–data comparison and delivering a system of analysis and visualization tools to the TCIS. The objectives and deliverables are:

•   Use the state-of-the-art Weather Research and Forecasting (WRF) model to generate ensemble simulations and deliver them to the TCIS.

•   Develop strategies for model–data comparison and deliver a system of analysis and visualization tools to the TCIS.

•   Investigate the impact of various physical parameterizations; compare to observations to determine the set of physical assumptions that produce simulated storms closest to the observed; suggest avenues to improve the representations of physical processes in the model and provide guidance to the research and operational communities.

B.   APPROACH AND RESULTS

1.   Investigating model sensitivity to the parameterization of physical processes

We have conducted WRF simulations to investigate the impact of cloud microphysics on hurricane track, structure, and intensity.

a.   Idealized simulations: In previous work [1,2], we developed a "real-ideal" WRF hybrid that has permitted us to conduct highly idealized hurricane simulations within the real-data WRF framework. This allows us to share the dynamical and physical framework used in realistic and operational simulations, thereby facilitating the transfer of knowledge gained through careful, controlled simulations to the operational WRF. Our previous work established that cloud microphysics materially influences the track of model tropical cyclones over periods as short as 1 to 2 days. Our present work has revealed, much to our surprise, that this previously established motion may emerge primarily through cloud-radiative feedback (CRF), since track variations nearly vanish when CRF is neglected. This study strongly motivates further research with respect to how condensation particles influence radiative processes and thus storm dynamics and thermodynamics [A].

b.   Real-case studies: The microphysical processes and their representation in hurricane models are of crucial importance for accurately simulating hurricane intensity and evolution. To investigate their impact, we compared and contrasted the members of high-resolution ensemble WRF model simulations of Hurricane Rita (2005) [B]. The members of the ensemble include simulations with two different sets of initial conditions, three different microphysical schemes, and eight different particle size distribution (PSD) assumptions within one of the microphysical schemes.

2.   Developing a strategy for model evaluation

In this project, we seek to identify strategies for comparing satellite data with model simulations. We aim at comparing statistical properties of the observed and simulated hurricanes, rather than point-by-point match. We have examined plots of storm track and intensity (Figure 2, left panel); time series (Figure 2, right panel); horizontal distributions of precipitation and clouds (Figure 3, left panel) exhibiting the storm extent and intensity variations that could

influence the storm tracks primarily due to the "beta effect" [1]; azimuthal averages of the cloud and precipitation parameters as a function of the distance from the center of the hurricane (Figure 3, right panel), to characterize the hurricane cloud structure; and vertical distribution of the azimuthal averages. Using the WRF-produced geophysical fields as input to forward models, we simulated radar and radiometer microwave observables. Examples include maximum radar reflectivity in the column (not shown), and brightness temperatures at various frequencies and polarizations (Figure 4). We compared the distributions of observed and modeled variables to provide richer information than comparing time and area means only. For example, the contoured frequency by altitude diagram (CFAD) displays the probability density function of storm characteristics (such as radar reflectivity) as a function of height (e.g., Figure 5, left column). It proves to be useful to describe storm structures and facilitate the comparison with satellite or in situ measurements [3]. To better describe the storm structure and to understand the interplay between the different geophysical parameters we have produced and analyzed 2-D joint distributions for various storm properties (e.g., Figure 5, right column).

Such comparison strategy centers on the storm-relative coordinate and avoids the exact geographical location match of actual and simulated storms, thus facilitates analyzing the model physics. Our analysis tools and results are now provided for incorporation in the JPL hurricane portal for public use.

3.   Results from evaluating the different parameterization schemes and suggested avenues from improvements in the representation of the physical processes

We performed a detailed analysis of our idealized simulations conducted with and without CRF. Figure 6 presents tracks from the CRF-on and CRF-off experiments. Six microphysics schemes were considered. In the CRF-on experiment, a fan of tracks similar to that obtained by [2] was obtained. The CRF-off tracks, however, exhibited very little motion diversity. The inset in Figure 6(b) reveals that CRF strongly impacts the winds of the hurricane outer core that control track as well as the radial distribution of diabatic heating, which likely controls those winds. This work strongly motivates further research not only on linking microphysics and radiation schemes, which tends to be very crudely handled at present, but also on comparisons with observations. Knowledge of radial and vertical distributions of diabatic heating in actual hurricanes is probably the clue to understanding why different models yield different tracks, and will likely lead to further material improvement in hurricane track forecast skill [A].

We incorporated into WRF a new radiation module, referred to as the Fu-Liou-Gu scheme, an improved version of the Fu-Liou scheme. Real-time WRF simulations for cirrus cases observed in the coastal and western United States have been conducted and compared with observations from MODIS and GOES-IR images. The new radiation module has been shown to work well in the WRF model and reproduce cirrus cloud distributions from satellite observations [C].

We used satellite observations from MODIS, TRMM-TMI, TRMM-PR, and QuikSCAT and compared them to the ensemble of WRF-simulated geophysical fields and their forward-computed microwave observables to determine those parameterizations (ensemble members) that yield a realistic forecast and those parameterizations that do not. Our results indi-

cate that multiparameter satellite observations can help discriminate between simulations with different microphysical assumptions [D]. In particular, assuming hydrometeor distributions with larger number of smaller particles results in model simulations with radiometric signatures that compare more closely to observations.

C.   SIGNIFICANCE OF RESULTS

Monitoring and understanding high-impact weather and climate change is one of the top priorities of JPL missions. The newly developed JPL Tropical Cyclone Information System addresses the need to connect satellite and in situ datasets with state-of-the-art numerical models to improve hurricane monitoring and forecasts. This study is directly aligned with the strategic goals of the TCIS development by augmenting the data portal with ensemble WRF simulations and model–data comparison tools and results.

The system of analysis tools developed by this project will be implemented in the TCIS during the next several months under a NASA-funded effort. The same NASA grant will support the TCIS incorporation of an operational high-resolution model forecast (Hurricane WRF–HWRF) during the 2010 hurricane season. These combined efforts will provide in the future vast comparisons of model simulations and satellite and in situ observations to allow comprehensive evaluation of model performance under different large-scale conditions. Based on the extensive and detailed evaluations, optimal physical options will be identified for tropical cyclone forecast.

This study demonstrates a pathway for effectively integrating JPL satellite data into university-led atmospheric science applications (e.g., modeling and diagnostic studies). The analysis tools developed here will become a critical component of the TCIS, helping to achieve the goals of the interagency Hurricane Forecasting Improvement Project (HFIP). As such, TCIS can potentially make a breakthrough contribution to hurricane forecast accuracy. This will increase the science value of JPL's, NASA's, and NOAA's satellite data by helping address an important scientific and societal problem, thus advocating JPL's current and future Earth science missions.

The model–data comparison strategies developed in this study will be useful for other model-data comparison projects pertaining to subjects such as climate change and global air quality. This study will also help to identify key atmospheric observations needed from new JPL missions for studying and improving hurricane forecasting, such as those recommended by the NRC Decadal Survey (e.g., PATH, ACE, GACM, etc.).

D.   FINANCIAL STATUS

The total funding for the JPL task was $99,630, all of which has been expended.

E.   ACKNOWLEDGEMENTS

We would like to thank the members of our JPL team who have worked hard to support our efforts. In particular, our thanks go out to P. Peggy Li (387) and F. Joseph Turk (334) for their continuous efforts to enrich our database with data from new instruments and missions and to include model data in the TCIS; Brian Knosp (328) and Quoc Vu (316) for their dedication and creativity in designing the satellite and model database and webtools; William Poulsen (334) for his efforts in providing the most accurate and up-to-date information on hurricane tracks; Van

Dang (324) and Tsae-Pyng Shen (383) for their efforts in analyzing MODIS and AIRS data and their comparison to observations; and Alexandra Chau (334) for her support in simulating Quik-SCAT observations from the model fields and developing the statistical comparisons to the QuikSCAT observations. Our sincerest thanks go out to them!

F.   PUBLICATIONS AND PRESENTATIONS

[A]   Robert G. Fovell, K. L. Corbosiero, A. Seifert, and K.-N. Liou, "Impact of Cloud-Radiative Processes on Hurricane Track," *Geophysical Research Letters*, in press.

[B]   Svetla M. Hristova-Veleva, Y. Chao, A. Chau, Z. Haddad, B. Knosp, B. Lambrigtsen, P. Li, J. Martin, W. L. Poulsen, E. Rodriguez, B. Stiles, J. Turk, and Q. Vu, "Using Instrument Simulators and a Satellite Database to Evaluate Microphysical Assumptions in High-Resolution Simulations of Hurricane Rita," AGU Fall Meeting, San Francisco, California, December 14–18, 2009.

[C]   Y. Gu, K. N. Liou, S. C. Ou, and R. Fovell, "Cirrus Cloud Simulations Using WRF with Improved Radiation Parameterization and Increased Vertical Resolution," to be submitted to *Geophysical Research Letters*.

[D]   Svetla M. Hristova-Veleva, H. Su, Y. Chao, A. Chau, Z. Haddad, B. Knosp, B. Lambrigtsen, P. Li, W. L. Poulsen, E. Rodriguez, B. Stiles, S. Tanelli, J. Turk, D. Vane, and Q. Vu, "Impact of Microphysical Assumptions on the Intensity and the Structure of Simulated Hurricanes: Can Satellite Observations Help Determine the Optimal Set of Microphysical Assumptions?" submitted to the AMS 29th Conference on Hurricanes and Tropical Meteorology, Tucson, Arizona, 2010.

G.   REFERENCES

[1]   Robert G. Fovell and Hui Su, "Impact of Cloud Microphysics on Hurricane Track Forecasts," *Geophys. Res. Lett*. **34** (2007): L24810. doi:10.1029/2007GL031723

[2]   Robert G. Fovell, Kristen L. Corbosiero, and Hung-Chi Kuo, "Cloud Microphysics Impact on Hurricane Track as Revealed in Idealized Experiments, *J. Atmos. Sci*. **66** (2009): pp. 1764–1778.

[3]   Sandra Yuter and Robert Houze, "Three-Dimensional Kinematic and Microphysical Evolution of Florida Cumulonimbus, Part II: Frequency Distributions of Vertical Velocity, Reflectivity, and Differential Reflectivity," *Mon. Wea. Rev*. **123** (1995): pp. 1941–1963.

H.   FIGURES
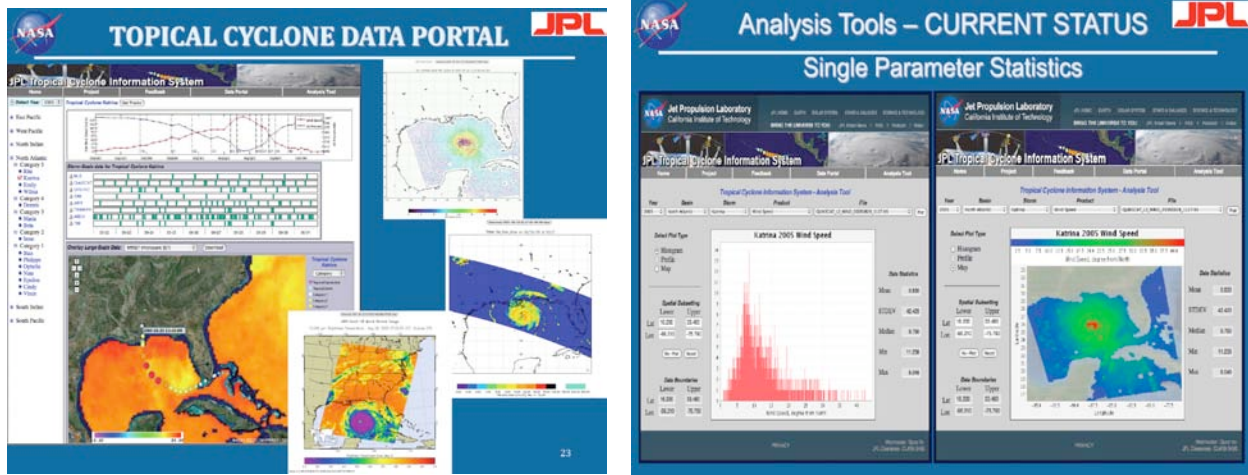
Figure 1. Illustration of the TCIS database (left panel) and the Analysis Tools (right panel).
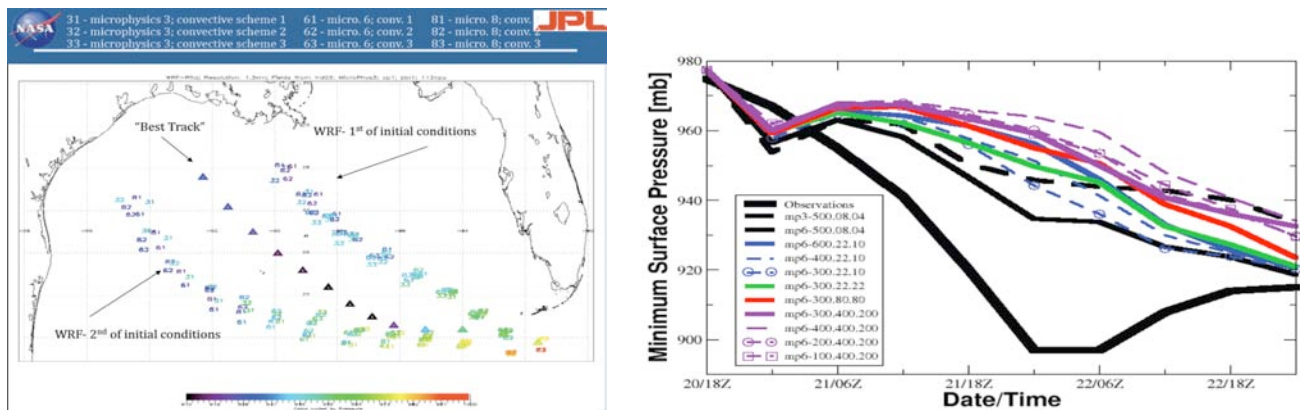
Figure 2. Examples of model–data comparisons in terms of storm track and intensity (left panel) and time series of minimum sea-level pressure (SLP, right panel). The two sets of simulated tracks resulted from using two different boundary conditions. The variations in the tracks and intensity within each set reflect the impact of the microphysical and convective parameterizations used by the model. The simulated storm positions in the left panel are color-coded as a function of the minimum SLP, reflecting storm intensity.
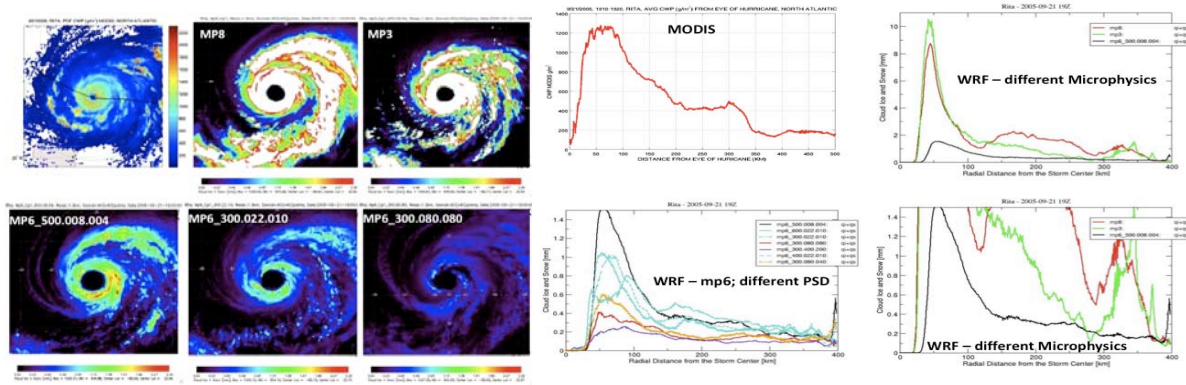
Figure 3. Comparison of the Ice Water Path observed by MODIS (top left) to the simulated by WRF cloud ice and snow fields. Two different microphysics and three different PSDs were used (the five graphs in the left panel). The right panel shows the azimuthally averaged cloud structures as a function of distance from the storm center. It is obvious that two of the microphysics (mp3 and mp8) produced structures that were very different from the observed. The naming convention is as follows: MPn_XXX.YYY.ZZZ where n signifies one of the WRF microphysics, XXX represents the assumed graupel density, YYY stands for the assumed intercept parameter of the rain PSD, and ZZZ for the assumed intercept parameter of the graupel PSD.
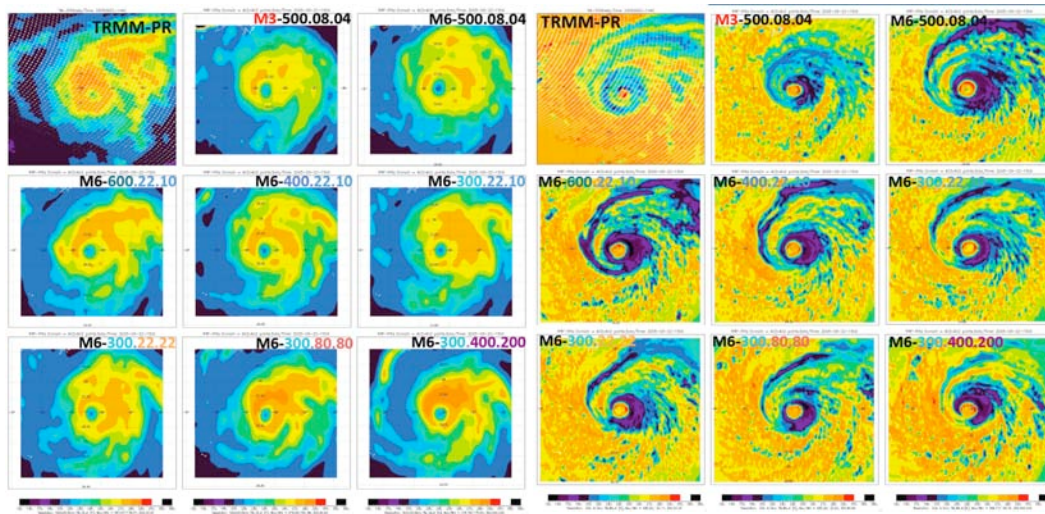
Figure 4. Microwave brightness temperatures at 19 GHz (left panel) and 85 GHz (right panel). The naming convention for the model simulations is as in Figure 3. The TRMM-observed brightness temperatures are in the top right corner of each panel. It is apparent that satellite observations can help discriminate between simulations using different microphysical assumptions.
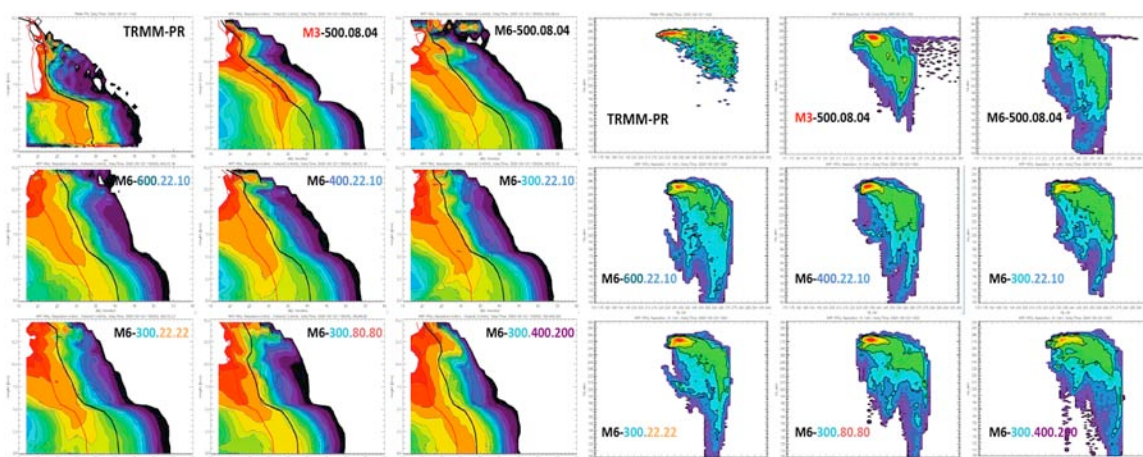
Figure 5. Examples of statistical comparisons. The left panel compares observed (top left) to simulated CFADs (contoured frequency by altitude diagrams [3]). The black profile represents the mean of the distribution while the red profile represents the mode. It is obvious that the model reflectivity has a broader distribution than the observed. Smaller PSDs result in reflectivity profiles that compare better to the TRMM observations. Right panel shows the joint distributions of the observed and modeled brightness temperatures at 19 and 85 GHz horizontal polarization. The naming convention for the model simulations is the same as in Figure 3.
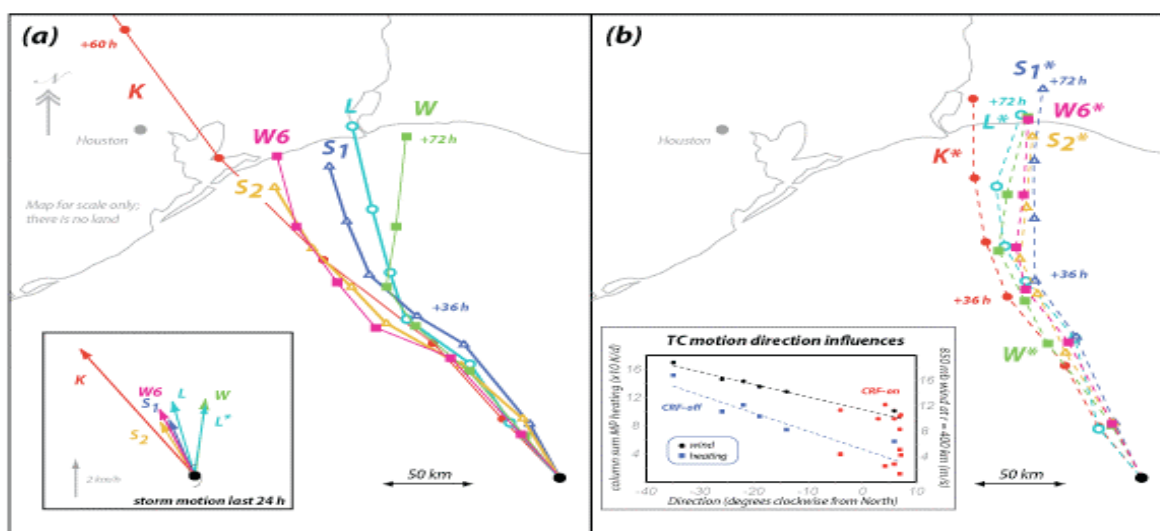
Figure 6. Tracks from idealized hurricane simulations testing the impact of the CRF — CRF-on and CRF-off experiments. See text for details.

ESTIMATING SOURCES OF ARCTIC MOISTURE USING GROUND- AND SPACE-BASED MEASUREMENTS OF WATER VAPOR AND ITS ISOTOPES

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1336

John Worden (PI), Earth Atmospheric Science (328)
Eric Posmentier, Department of Earth Sciences, Dartmouth College
Xiahong Feng, Department of Earth Sciences, Dartmouth College

A.   OBJECTIVES

New techniques of inverting data from satellite-borne radiometers to obtain atmospheric profiles of HDO ($^1H^2H^{16}O$ water) vapor concentration, as well as the profiles of other trace gas concentrations, creates nascent possibilities for studying atmospheric structure, processes, and dynamics (e.g., [1,2]). However, the quantitative value of remotely sensed HDO has not yet been determined. To help evaluate such data, we begin by posing the question: What is the potential information content of the deuterium isotope ratio of atmospheric water vapor as a constraint on atmospheric variables and processes? We also pose the corollary question: What is the incremental value, over conventional meteorological and climatologic observations?

A primary deliverable from this SURP proposal is a numerical model capable of simulating the distribution of atmospheric HDO vapor under differing temperatures, winds, and mixing rates. Just such a model was developed for the distribution of another heavy isotope of water, $H_2^{18}O$, and used to evaluate seasonal changes in atmospheric vertical mixing rates in the boundary layer [3]. For this study, the same model was expanded to include HDO in addition to $H_2^{18}O$. This model is 2-D vertical-horizontal. In this study, the horizontal coordinate is longitude, so the model plane is a zonal section. Surface and atmospheric temperatures are specified. Zonal wind speeds and the zonal and vertical mixing rates are also specified. Equations representing conservation of mass for normal water ($H_2^{16}O$) and the two heavy isotopes (HDO and $H_2^{18}O$) are solved, and the rates of change of all three isotopes are integrated until a steady state is obtained.

The model geometry includes 23 pressure levels from the surface up to 150 hPa. There are 20 segments in the horizontal direction, but the horizontal grid size is not constant. This scheme makes better resolution possible in the high-gradient coastal regions while maximizing overall computational speed. The model contains modules for the rate of surface evaporation and the precipitation rate, using different parameters for the three isotopes, thus implicitly specifying both the kinetic and equilibrium fractionation factors. The dependence of fractionation on temperature and air pressure is both provided for. At this time, the model does not include re-evaporation of falling hydrometeors, but this process will be added to the next version.

The numerical experiments performed to evaluate the information content of DHO all involve a pair of runs in which a single specified variable, or a single specified process-related parameter, was assigned each of two values. A large difference between the HDO values in the pair implies that HDO measurements is sensitive to, and can be used to constrain, the variable or parameter that was changed. Inversely, a small difference implies that HDO is too insensitive to be a useful constraint on the variable or parameter being investigated. In these interpretations, large and small are relative the minimum difference resolvable by TES (Tropospheric Emission Spectrometer) methodology. In cases where more than one variable or parameter can cause the same change in HDO, there is an ambiguity typical of those that frequently arise in the well-known "geophysical inverse problem." We comment further on this later.

While our original goal was to use this tool for constraining estimates of Arctic evaporation strength, we realized during this work that this objective was impractical due to the lack of evenly distributed TES data under different cloud regimes at high latitudes. However, equally important objectives became readily apparent and more achievable upon further consideration, which we describe next.

B.   APPROACH AND RESULTS

HDO as an estimator of monsoon circulation strength — the first experiment (detailed in this report) deals with the monsoon (e.g., Walker) circulation. This is the seasonally reversing flow of air from the oceans onto the continents in the summer, rising, diverging aloft and flowing out to the oceans, and descending. It is part of the monsoon system, and contributes to the higher precipitation over continents and drier conditions over the ocean in the summer. It is directly related to the summer continental low-pressure systems and marine highs. Inverse conditions prevail in the winter.

The Walker circulation, as mentioned, reverses from summer to winter, and its summer and winter speeds vary interannually. The latter variability is linked to widespread flooding and droughts. However, these speeds are ambiguous in definition and virtually impossible to measure directly, causing most monsoon studies to be lacking in essential data. So this first experiment focuses on the effect of the Walker circulation speed on the distribution of HDO, and the possibility of using the HDO as a proxy for Walker circulation.

Figure 1 shows a zonal section of the change in δD (the anomaly of the HDO–H_2O ratio, per mil) resulting from reversing the Walker circulation at a speed of 1 ms^{-1}. The changes in δD are large enough to be resolved by TES at all altitudes above 1 km over the continent, and reach a very strong maximum of about 100 per mil in the upper troposphere at about 8 km. Over the ocean, the changes are discernible only in above the tropopause, but they do become quite large — ca. 70 per mil) in the lower stratosphere. Even in situ measurements of δD are virtually impossible here because of the extreme dryness at these altitudes.

C.   SIGNIFICANCE OF RESULTS

To summarize, this experimental pair strongly suggests that the strength of the Walker circulation — a component of the monsoons — can be inferred from TES observations of δD.

However, is it possible to measure this circulation by conventional methods? With today's anemometry technology, it is virtually impossible to measure a profile of convergence around the margins of a continent. Even the precipitation caused by convergence is difficult to measure accurately, especially in polar regions and deserts, and spatial resolution is insufficient in heterogeneous regions where monsoon precipitation is often strongest.

Perhaps it's possible to use vapor mixing ratio (q) measurements from TES or other missions as proxies for the Walker circulation strength. Vapor mixing ratio is another diagnostic of the model, so we were able to test this possibility directly. Unfortunately, the model results show that q is sensitive to Walker circulation only in the lowest 2 km over land (Figure 2). The value of q in this air is subject to extremely large spatial and temporal "noise" resulting from the diel cycle, shifting synoptic patterns, and surface heterogeneity, so q would not be adequate as a Walker circulation proxy. TES may therefore be the only hope for obtaining a meaningful proxy for monsoon strength.

Summary and Conclusions

A.   The atmospheric hydrology/isotope model simulates realistic distributions of vapor isotopes.

B.   The sensitivities of vapor deuterium ratios to several atmospheric variables and processes are large relative to the resolution of TES observations.

C.   Deuterium responds to both global- and synoptic-scale perturbations.

D.   The informational content of deuterium is not redundant with water vapor concentration.

D.   NEW TECHNOLOGY

The modeling capability delivered to JPL will allow us to test different measurement strategies for constraining important estimates of important components of the tropical and mid-latitude water cycle.

E.   FINANCIAL STATUS

The total funding for this task was $100,000, of which $90,143 has been spent.

F.   PUBLICATIONS

None.

G.   REFERENCES

[1]   John Worden, David Noone, Kevin Bowman, et al., "Importance of Rain Evaporation and Continental Convection in the Tropical Water Cycle," *Nature* **445** (February 1, 2007): pp. 528–532.

[2]   Christian Frankenberg, et al., "Dynamic Processes Governing Lower-Tropospheric HDO/H_2O Ratios as Observed from Space and Ground," *Science* **325** (September 11, 2009): pp. 1374-–1377.

[3]   Eric S. Posmentier, et al., "Seasonal Variations of Precipitation $\delta^{18}O$ in Eastern Asia," *Journal of Geophysical Research* **109** (2004): D23106). doi:10.1029/2004JD004510

H.   FIGURES
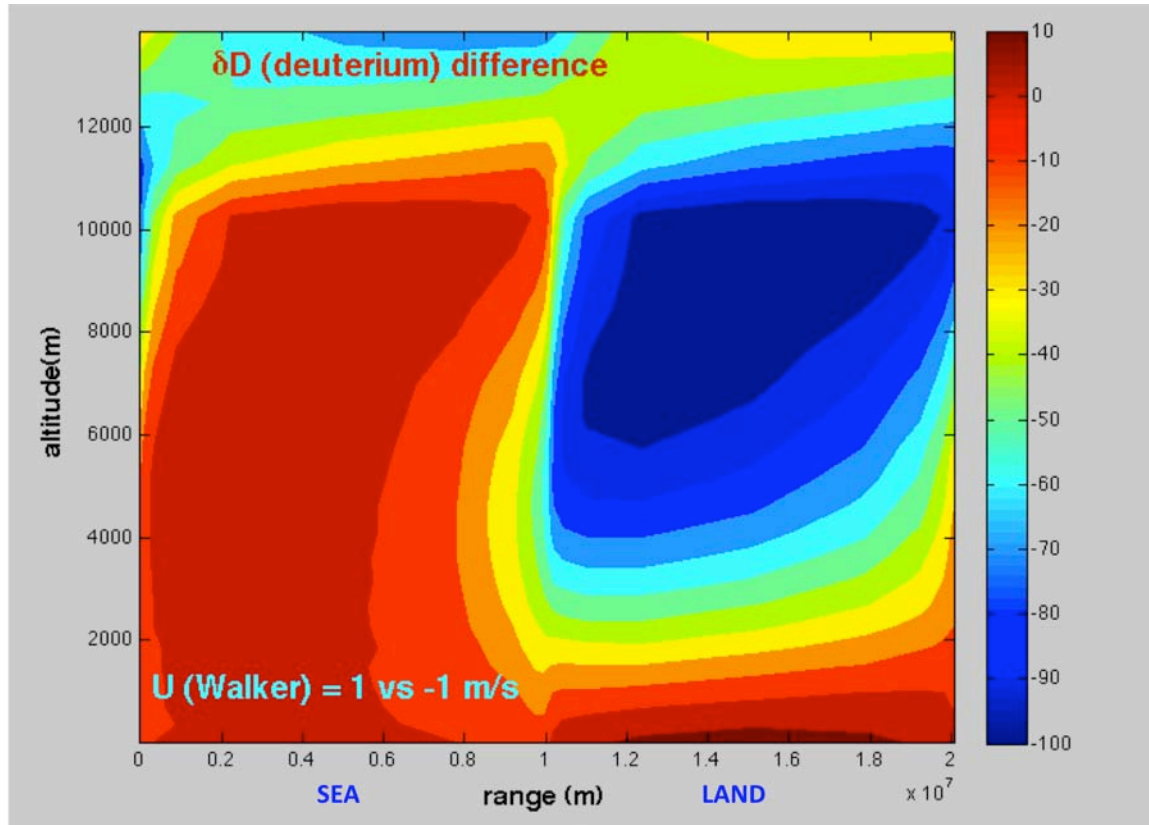
Figure 1. Zonal section of the change in δD (per mil) resulting from changing the Walker circulation speed from +1 to −1 ms^{-1}. The range interval from 0 to 1×10^7 m is ocean; 1 to 2×10^7 m is land.
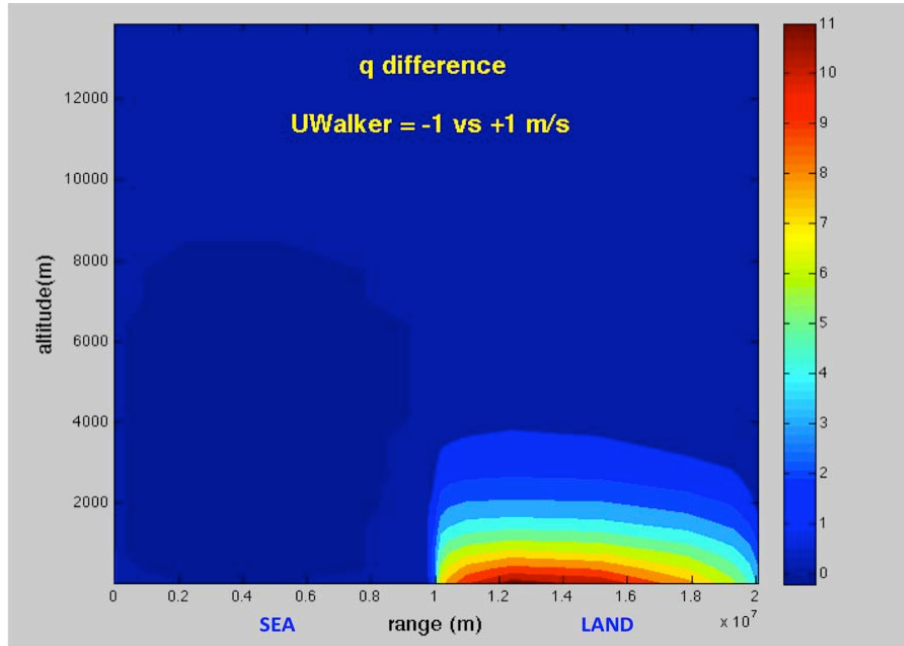
Figure 2. Zonal section of the change in q (g/kg) resulting from changing the Walker circulation speed from +1 to −1 ms^{-1}.

**ENABLING AUTONOMOUS HUMAN MISSIONS
TO THE MOON AND MARS**

ERROR-CORRECTING CODES FOR RADIATION-TOLERANT MEMORY

Michael K. Cheng (PI), Communication Architectures and Research Section (332)
Vijayakumar Bhagavatula (Co-PI), Department of Electrical and
Computer Engineering, Carnegie Mellon University
Seungjune Jeon, Department of Electrical and
Computer Engineering, Carnegie Mellon University
Euiseok Hwang, Department of Electrical and
Computer Engineering, Carnegie Mellon University

A.   OBJECTIVES

The objective of this SURP research effort was to develop new error-correction strategies that would produce "radiation-hardened" reliability performance and achieve storage efficiency on a standard commercial-off-the-shelf (COTS) memory device. For space applications, memories are either made radiation-hard by strong material shielding or made radiation-tolerant through protection of error-correction codes (ECC). However, the ECC scheme used by recent JPL missions does not handle multiple bit errors at all. For example, the Mars Exploration Rovers (MER) adopted a single-error-correction and double-error-detection (SECDED) code to protect memory devices against radiation-induced faults. This simple error detection and correction (EDAC) code cannot sustain an event with three or more bit errors and requires a high overhead for redundancy. Even so, this SECDED code remains as a baseline in data storage protection for upcoming JPL missions.

In this work, we devised a channel that models radiation effects on memories in space. We compared the performance of a modern ECC that can take advantage of the side information offered by the channel to that of classical ECCs where side information is not used. We show through Monte Carlo simulations that the memory device life can be extended significantly by using our new approach over the conventional SECDED scheme, especially when the fundamental application storage units are kilobytes or larger.

B.   APPROACH AND RESULTS

Digital data are stored on memory devices as bits. To recover bits that might undergo a flip in polarity (changing from a "0" to a "1" or vice versa), we can apply a "block-based" ECC scheme where we group a collection of bits into an information block and map this information block to a ECC codeword with redundancy. The structure of the ECC defines the relationship between the redundant bits and the information bits. The ratio of the information bits per ECC

codeword length is the code rate and indicates the efficiency of the code. For memory applications, a strong ECC with a high code rate is desired.

The baseline ECC for memory protection in previous and current JPL missions is the (39,32) Hamming code where every 32 information bits are mapped to a 39-bit codeword and this code has a code rate of 32/39. However, this ECC was developed many decades ago when the memory device storage capacity was small (Mbits or less). As memory capacities increase and applications treat information storage in larger chunk sizes, the small SECDED (39,32) Hamming code no longer offers sufficient error-correction capability. We can correct more errors and still increase the code rate when compared to Hamming codes by using longer and more modern ECCs. The focus of our work was on low-density parity-check (LDPC) codes, but we also considered Bose-Chaudhuri-Hocquenghem (BCH) codes and Reed-Solomon (RS) codes. We compared the performance of these codes against the baseline and showed that LDPC codes, together with side information provided by radiation parameters, can increase the life span of COTS memory devices in space. Our research followed these steps in sequence:

1.   <u>Derived a probabilistic model of memory failures in a radiation environment</u>. There are two types of errors in memory systems: soft errors (temporary) and hard errors (permanent). Soft errors may be caused by energetic particles, coupling from power supply noise, or variability in device behavior. Typically, soft errors would produce only a single cell malfunction. As memory building blocks shrink into the nanometer regime and when memories are used in space applications, the frequency of multiple cell malfunction increases and these error events can range from a few bit flips to hundreds of errors. In addition to soft errors, hard errors can occur both in manufacturing due to defects or lithography contaminants and in use due to device wear-out or cosmic radiation.

We devised a simple channel model that accounts for both soft and hard errors, as seen in Figure 1, where T is the time interval of radiation exposure. This channel model assumes that bit errors are independent and identically distributed. The probability of a bit "0" flipping to a bit "1" (or soft error) in time interval T is given by $p_{01}(T)$. The probability of a bit "0" becoming a stuck bit (or hard error), marked by ε in Figure 1, is $q_0(T)$. The probability that a bit "0" remains error free is $1 - p_{01}(T) - q_0(T)$. The actual bit error probabilities, $p_{01}(T)$ and $q_0(T)$, were derived by solving a set of differential equations. (For details, please refer to [C] in the publication list.) These equations were obtained based on the constraint that the error probabilities over time $T_1 + T_2$ must have some equivalence relationship to individual error probabilities calculated in each time interval T_1 and time interval T_2. For a given soft error rate λ and hard error rate λ_e, we plotted the various bit error probabilities in Figure 2.

Based on this model, we can treat each bit stored in the memory device as a likelihood that a bit position is a "0" or a "1" instead of strictly a "0" or "1" with certainty. The structure of LDPC codes allows us to decode a codeword by iterating and refining the bit likelihoods until all of the bit decisions agree and converge to a codeword. Or the decoding stops when the maximum number of iterations is reached and a decoding failure declared. This iterative decoding approach enables better coding gains than the conventional decoding of Hamming, BCH, and RS codes where each code symbol is treated strictly as a "0" or "1" in decoding.

2.   <u>Evaluated performance of selected error correcting codes on the derived channel model</u>. We compared the RS, BCH, and LDPC decoding performance on the derived model of Figure 1. The code parameters (information length and code rate) for each code were selected to be as close as the design allowed ensuring a fair comparison. We analytically calculated [1] the decoding failure probability of RS and BCH codes. However, no theoretical methods are available to compute the exact probability of decoding failure for LDPC codes. We evaluated the performance of LDPC codes by software Monte Carlo simulations. Since the software decoder throughput is limited, we could not obtain very low codeword error rates for LDPC codes. The error events were simply too rare to be observed in a reasonable amount of simulation time.

Scrubbing is an operation to refresh memory content with the ECC decoder output to correct errors. The scrubbing interval is the time between each scrubbing. A scrubbing interval of one hour is assumed in this report. The number of soft errors can be reduced after scrubbing but the number of hard errors will only accumulate regardless of scrubbing.

We plotted the codeword error rate for each RS, BCH, and LDPC code versus time. Each time unit is one hour, so 2000 hours corresponds to approximately 83 days. The parameters for the codes are all selected to be as close as possible so that the comparisons are fair. We had to choose a relatively high soft and hard error rate $\lambda = \lambda_e = 10^{-3}$ in order to make our simulation time reasonable. Our selection is representative of an overly stringent radiation environment. The radiation-induced soft and hard error rates are generally less than 10^{-6} for most missions. But at this low error rate, software simulations would require months to produce any result.

a.   Figure 3 illustrates that the codeword error rate of the 4096-bit LDPC code (deigned by Progressive Edge Growth [2] method) grows much slower than that of equivalent BCH or RS codes. Therefore, LDPC codes can extend the life span of memory systems in a radiation environment. For example, at a target codeword error rate of 10^{-4}, the memory system using LDPC code can operate up to 1300 hours, whereas the memory system with BCH code can operate only up to 225 hours. The worst performance is given by RS code at 175 hours.

b.   Figure 4 shows performance for codes with 1024-bit information length. We also included as a comparison the high rate AR4JA LDPC code designed at JPL [3]. And this code is part of the Consultative Committee for Space Data Systems (CCSDS) standard. Again, LDPC code outperforms RS and BCH codes.

By using a modern LDPC code and decoding the bit likelihoods provided by the channel, we could extend the memory life and therefore, mission length, by manyfold.

3.   <u>Computed the undetected codeword error rates</u>. An undetected codeword error denotes the event when the decoder produces a codeword that is not equal to the original codeword. This error is different than a decoding failure because, unlike the failure case where the decoder will flag the output as erroneous, the decoder in the undetected error case believes that its output is correct but in fact the decision is not what was originally stored in memory. The result is undesirable because the application that uses the decoded information has no knowledge that the information is incorrect and might use the data as if it were error-free. The undetected

error rate for Hamming codes is significant enough that it cannot be ignored, especially if the code is used to protect a large memory device where many codewords are used. We analytically computed the undetected codeword error rate for the (39,32) Hamming code that has been used in previous JPL missions.

Binary Hamming codes are constructed such that the codeword length is $2^m - 1$ and the information length is $2^m - m - 1$, where m is an integer greater than zero. Also by construction, Hamming codes always have a minimum distance (d_{min}) of three. The minimum distance is the least number of bit positions any two codewords in a code differ by. Any binary block code can correct up to $\left\lfloor \dfrac{d_{min} - 1}{2} \right\rfloor$ bit errors. A Hamming code is a perfect code with $d_{min} = 3$. Therefore, Hamming codes are single error correcting and cannot detect double errors. Note that we can only construct a (63,57) or a (127,120) binary Hamming code. The (39,32) binary Hamming code is actually derived from the (63,57) Hamming code by first not using all of the odd weight codewords (or codewords with an odd number of ones). Since half of the codewords in the code are thrown out and not used, the code dimension is halved and we have the (63,56) Hamming code. We also increased the minimum distance to $d_{min} = 4$ by doing this expurgation. Increasing the minimum distance by one allows this code to detect double errors also and decreases the undetected error rate significantly. Now, we shorten the code by not using 24 of the information bits and arrive at the (39,32) binary Hamming code.

We plotted in Figure 5 the probability of undetected codeword error of binary Hamming codes with different lengths versus the raw bit error rate. The undetected error probability of the (39,32) Hamming code is larger than 10^{-5} when the raw bit error rate is higher than 10^{-2}. For a typical memory device that has a capacity of 128 Gb, which maps to about 3.2×10^9 (39,32) Hamming codewords, there would be more than 32000 undetected codeword errors if the raw bit error rate on the memory device were 10^{-2} or higher. Small Hamming codes are not strong enough to protect large devices for a long period of time and the undetected codeword error rate will increase, with deteriorating device bit error rate, over time to a non-negligible value.

The undetected error rates for BCH and LDPC codes are much lower than Hamming codes due to their larger minimum distances. The actual undetected error rate for each code is difficult to obtain analytically. The probability of miscorrections for either code is so low that we could not even find one event in a weeklong simulation.

4.   Calculated error rates on a current COTS memory device. We analyzed the performance of using LDPC, BCH, or Hamming code on a COTS NAND flash memory device that is under consideration for upcoming JPL missions. This memory architecture has 4314 bytes per page. The fundamental storage unit in bytes depends on the specific application and could easily be page-based or larger, especially when many potential uses such as high-fidelity images constitute large sizes. In these cases, a longer code, especially one that matches to the fundamental storage unit, can achieve much better error protection.

Manufacturing COTS memory is not a perfect process either. There are inherent defects on a device, but as long as these defect counts are small, using an appropriate ECC to

protect stored data can overcome these device imperfections. For the 32- to 128-Gb devices that we considered, the manufacturer suggests an ECC that can correct up to 12-bit errors for every 539 bytes of data. We therefore, compared the error rate performance of these devices when protected by the (39,32) Hamming code, the (4252,4096) BCH code, and the (4314,4096) LDPC code. The minimum distances of the Hamming code and the BCH code are known and the codes can correct single- and 12-bit errors, respectively. The minimum distance of the LDPC code is not known and we obtain the performance measure through Monte Carlo simulation.

Assume that the user treats the smallest storage unit as a page or 4314 bytes. Figure 6 compares the page error rates for Hamming, BCH, and LDPC codes. A page error is declared as long as one or more codewords in the page cannot be decoded. A page error rate of 1 (indicates that memory is no longer useful) is reached in only 50 hours when Hamming code is used. BCH code extends the memory life to 100 hours because a page spans fewer BCH codewords than Hamming codewords. LDPC code extends the memory availability to more than 700 hours due to the gain in iterative decoding of bit likelihoods rather than absolute 0's and 1's.

C.   SIGNIFICANCE OF RESULTS

A recent and exciting development in the area of error-correcting codes on the additive white Gaussian (AWGN) channel is the design of codes and the corresponding iterative decoding algorithms that allow the decoder to repeatedly refine the bit likelihoods of a received (or memory stored) word until convergence to a codeword. This combination of code design and low-complexity decoding results in performance close to the Shannon Capacity. In this work, we developed a channel model that captures radiation effects on storage systems, and this model enables application of these new codes and iterative decoding algorithms to COTS memory devices. We demonstrated that LDPC codes could outperform BCH, RS, and Hamming codes in a radiation environment so that the life span of memory systems can be extended in space. Our probabilistic channel model is valid as long as the bit errors in memory systems are not correlated. Our work paves a possible path for JPL to achieve "radiation-hardened" error rate performance on a COTS memory device that otherwise is only attainable with a heavily shielded custom memory part.

D.   NEW TECHNOLOGY

Under this SURP research, we designed a channel model that captures the effects of radiation on memories in space. Using parameters that indicate the strength of radiation and the scrubbing period, we developed a new technique to generate symbol likelihood information for iterative decoding of LDPC codes on our channel. Through simulations, we showed that the use of LDPC codes for memory protection could increase the life span of COTS memory devices manyfold when compared to current schemes using Hamming codes. Not only does the error rate performance improve, the storage efficiency (percentage of device available for information storage) is also significantly increased.

E.   FINANCIAL STATUS

The total funding for this task was $100,000, of which $97,740 has been expended.

F. ACKNOWLEDGEMENTS

G. PUBLICATIONS AND PRESENTATIONS

[A] Seungjune Jeon, Euiseok Hwang, B.V.K. Vijaya Kumar, and Michael K. Cheng, "Investigation of Memory Protection Using LDPC codes," 10th Annual Non-Volatile Memory Technology Symposium (NVMTS'09), Portland, Oregon, October 2009.

[B] Seungjune Jeon, Euiseok Hwang, B.V.K. Vijaya Kumar, and Michael K. Cheng, "LDPC Codes for Memory Systems with Scrubbing," submitted to IEEE Global Communications Conference 2010 (GLOBECOM 2010), Miami, Florida, 2010.

[C] Seungjune Jeon, B.V.K. Vijaya Kumar, Euiseok Hwang, and Michael K. Cheng, "Evaluation of Error Correcting Codes for Radiation-Tolerant Memory," *The Interplanetary Network Progress Report* **42-181**, Jet Propulsion Laboratory, May 15, 2010.

H. REFERENCES

[1] G. C. Cardarilli, et al., "Data Integrity Evaluations of Reed Solomon codes for Storage Systems," in *Proc. 19th IEEE International Symposium on Defect and Fault Tolerance in VLSI Systems (DFT04),* Cannes, France, October 2004.

[2] Z. Li and B. V. K. Vijaya Kumar, "A Class of Good Quasi-Cyclic Low-Density Parity-Check Codes Based on Progressive Edge Growth Graph," in *Conf. Record of the 38th Asilomar Conf. on Sig., Sys, and Comp* **2**, Pacific Grove, California, November 2004, pp. 1990–1994.

[3] K. S. Andrews, D. Divsalar, S. Dolinar, J. Haminks, C. R. Jones, and F. Pollara, "The Development of Turbo and LDPC Codes for Deep-Space Applications," *Proc. IEEE* **95** (November 2007): pp. 2142–2156.

I. FIGURES
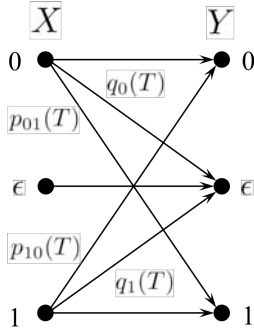
Figure 1. Channel model with radiation effects.

(a) Probabilities (b) Conditional probabilities
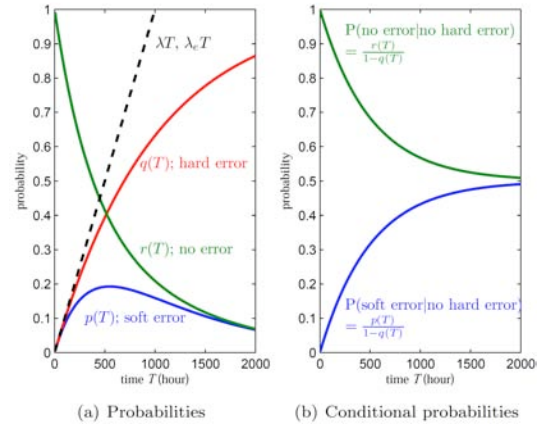
Figure 2. Stringent radiation environment $\lambda = \lambda_e = 10^{-3}$ errors/bit/day.
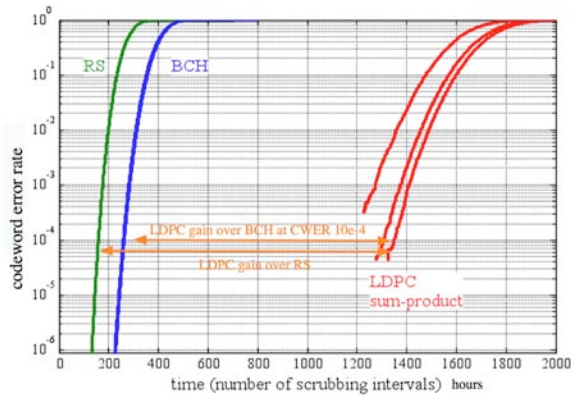
Figure 3. All codes are of information length 4K bits and code rate 0.89.
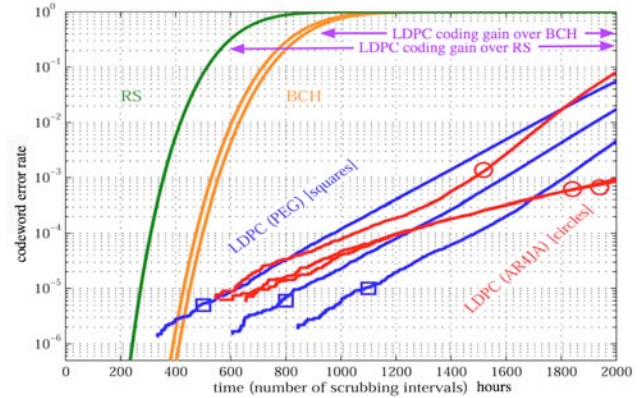
Figure 4. All codes are of information length 1K bits and code rate 0.89.
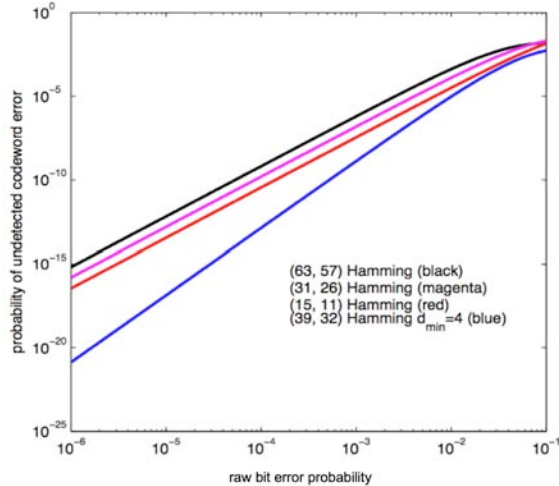
Figure 5. Undetected error probability of Hamming codes.
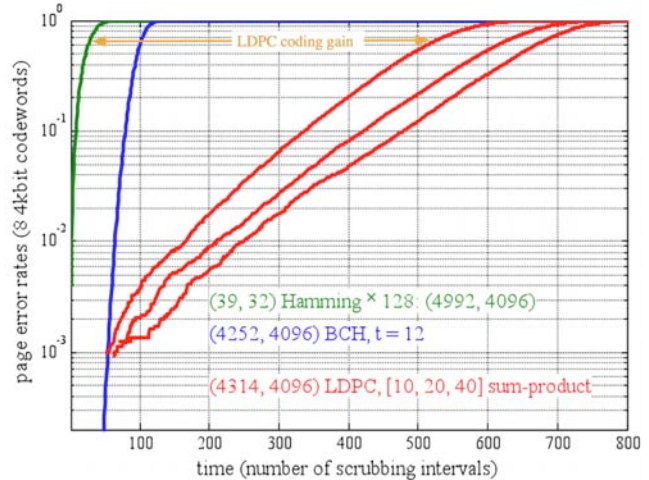
Figure 6. A page of 4314 bytes is 862 Hamming, 8 BCH, or 8 LDPC codewords.

JOINT SOURCE CHANNEL CODING FOR DEEP SPACE IMAGE TRANSMISSION

Dariush Divsalar (PI), Communications Architectures and Research Section (332)
Giuseppe Caire (Co-PI), Department of Electrical Engineering,
University of Southern California
Ozgun Bursalioglu Yilmaz, Department of Electrical Engineering,
University of Southern California

A.   OBJECTIVES

Shannon's source-channel separation principle states that, in the limit of large block length, the optimal performance can be approached by independently designing the source coding and the channel coding schemes [1]. Driven by the separation principle, source coding such as image compression and video coding, and channel error correcting coding, have been developed by different communities of researchers according to quite different viewpoints. Unfortunately, the concatenation of separately designed source and channel coding may be fragile to residual channel errors, due to the well-known catastrophic behavior of conventional source coding with respect to errors in the coded bit stream. In [2,3] we proposed a new approach to the efficient design of joint source-channel coding (JSCC) schemes that overcome this problem. The central novel idea of the proposed JSCC scheme consists of eliminating the catastrophic effect of residual bit error rate (BER) by mapping directly the redundant image transform coefficients onto channel-coded symbols via a linear encoder. Good linear encoders are known to be noncatastrophic and efficient modern forward error-correction (FEC) code design as proposed in [4] can be leveraged. The general objective of this task is to develop a novel general JSCC scheme for the transmission of high-resolution digital images over the noisy deep-space link by building on the approach of [2,3], and leveraging the powerful code design of [4].

B.   APPROACH AND RESULTS

1.   System Setup

We consider a source formed by s independent (or "parallel") components. A block of source symbols of length k is denoted by the $s \times k$ array S, where the i-th row of S, denoted by $S_{(i)}$, corresponds to a source component of the parallel source. The channel is modeled as a binary-input AWGN (BI-AWGN) channel, with energy per symbol Es and noise per real component equal to $N_0/2$. A JSCC scheme is formed by a mapping f from S to a codeword X of length N over the binary symbols $\{+1,-1\}$. We define the spectral efficiency $b = N/(sk)$ as the number of (channel coded) bits per source sample, in analogy with the conventional measure of efficiency in "bit per pixel" of image coding. Letting d_i denote the mean-square error

(MSE) distortion for the *i-th* parallel source component, we consider as the overall system performance measure the total weighted MSE (WMSE) distortion given by

$$D = \frac{1}{s} \sum_{i=1}^{s} v_i d_i$$

where $\{v_i : i = 1,..., s\}$, are weights that depend on the specific application. The optimization of a coding scheme in the case of ideal capacity achieving channel codes consists of allocating optimally the distortions d_i for each source component in order to solve the minimum WMSE problem:

minimize $\quad \frac{1}{s} \sum_{i=1}^{s} v_i d_i$

subject to $\quad R\left(\frac{1}{s} \sum_{i=1}^{s} v_i d_i \right) \le bC, \quad \sigma_i^2 \ge d_i \ge 0, \quad \forall i,$

where σ_i^2 denotes the variance of the *i-th* source component.

2.   Wavelet Transform and Quantization

Figure 1 (left) shows the subbands of a *W*-level discrete wavelet transform (DWT). In this task, we considered the CDF 9/7 [5] wavelet used by the state-of-the art JPEG2000 image coder. Figure 1 (right) shows the histogram of the DWT coefficients for different subbands, for the test MER1 image, corresponding to the file 1F178787358EFF5927P1219L0M1.pgm, kindly provided by Dr. A. Kiely of JPL. All subbands yield approximately symmetric non-Gaussian distributions, showing concentrated around zero. In contrast, the LL0 subband has a nonsymmetric distribution that is not quite suitable for direct quantization. Therefore, we applied a discrete cosine transform (DCT) to the LL0 subband, obtaining a nicely symmetric and concentrated distribution of the DCT coefficients with the exception of a few "outlier" coefficients (indicated by arrows in Figure 1b) that must be directly encoded in the header information. For quantization of the DWT coefficients, we considered a family of embedded dead-zone scalar quantizers as used in the JPEG2000, Part I [6]. In Figure 3 (left), the embedded dead-zone quantizer is shown for 3 different resolution levels. The parameter Δ is adjusted depending on the dynamic range of the given subband. At resolution level p, the quantizer has $2^{p+1}-1$ regions. We use the ternary alphabet $\{A,B,C\}$ to index these regions. For every subband, the operational rate-distortion function of embedded scalar quantization is given by the lower convex envelope of the corresponding entropy-distortion points Figure 3 (right) shows qualitatively the shape of the quantization operational rate-distortion function. Using these curves in the minimum WMSE problem, we obtain the optimal quantization bit allocation for the different source subbands. We checked that a near-optimal rate-distortion performance can be achieved by this scheme, in line with the ICER performance.

3.   Linear Index Coding

Conventional fixed-to-variable-length entropy coding, such as Huffman or arithmetic coding [6], is catastrophically nonrobust to residual channel errors. In order to cope with noisy channels, we investigated the use of channel-optimized quantization [7,8], also known as "index coding" or "fixed-length index assignment." Recently, [2,3] use linear coding in order to efficiently implement fixed-length index assignment. In this way, each quantizer symbol is directly linearly encoded into a channel codeword. The resulting scheme provides naturally a channel-optimized successive refinement quantization scheme, and therefore it is ideally suited to the problems considered here. Since the quantization layer symbols are ternary while we can design efficient binary codes, we define a binary encoding of the ternary symbols and encode separately two layers of coding per quantizer level. The rate of the code at a given layer is chosen according to the transmissibility condition

$$kH_p^{(i,u)} = n_{p,i,u}C/(1+\theta_{p,i,u})$$

where $H_p^{(i,u)}$ is the entropy of the binary upper sublayer of the *p-th* quantization layer of the *i-th* subband, and $\theta_{p,i,u}$ denotes the overhead (relative gap to capacity) of the corresponding binary linear code. At the decoder, we use iterative belief propagation (BP) decoding incorporating the a priori statistics of the quantization layers, and multistage decoding in order to decode the layers in sequence. We considered a specially designed family of rate-compatible punctured protograph low-density parity-check (LDPC) codes based on the protograph shown in Figure 6. Each node corresponds to 1024 nodes in the actual graph. Puncturing patterns in order to span rates from 1/3 to 16/17 have been optimized. EXIT chart analysis of the convergence of the BP decoder for BI-AWGN capacity corresponding to *SNR = Es/N$_0$ = 3* dB and for all empirical quantization index entropies was carried out, in order to arrive at a complete rate allocation table. Figure 4 shows the performance over the BI-AWGN channel. This performance is obtained in the infinite-length assumption and suffers from 2 dB of reconstruction SNR loss with respect to the conventional concatenation of ICER with a capacity-achieving code for the same BI-AWGN. Our goal for the future is to close this gap at infinite-length by further optimization of the code family and more accurate mapping of the source bits onto the variable nodes of the protograph, and finally run extensive simulations in finite length and finite complexity, in order to fully capitalize on the promised robustness of the scheme. Finally, we shall investigate the design of scalable schemes that provide variable quality of reconstruction at different operating channel SNRs. This is useful in the case where, due to fading or random antenna misalignment, the received SNR fluctuates unpredictably around its nominal value. The same convex optimization framework used before to solve the minimum WMSE problem can be extended to encompass a range of possible channel SNRs.

C.   SIGNIFICANCE OF RESULTS

The results achieved are promising but not yet fully satisfactory. At present, the comparison with the state-of-the-art ICER image coder concatenated with state-of-the-art channel codes still suffers from a 2-dB gap (about 0.4 bpp rate gap at channel Es/N_0 = 3dB). This gap is due to a not fully optimized protograph code family and to the fact that we haven't yet optimized the mapping of the source bits onto the variable nodes of the protograph. Nevertheless,

the scheme has potential to provide competitive results in terms of robustness to channel SNR mismatch. Further, we plan to develop a multiresolution/unequal error protection version of our scheme that will be able to span a wide range of channel SNRs.

D. NEW TECHNOLOGY

The new technology resulting from this project consists of a novel image coding over noisy channels scheme that potentially can replace the current state of the art. The new scheme has low complexity at the encoder side, and holds promise of an improved robustness to channel SNR mismatch.

E. FINANCIAL STATUS

The total funding for this task was $100,000, all of which has been expended.

F. ACKNOWLEDGEMENTS

G. REFERENCES

[1]   T. Cover and J. Thomas, *Elements of Information Theory*, 2nd Edition, Wiley, 2002.

[2]   M. Fresia and G. Caire, "Combined Error Protection and Compression with Turbo Codes for Image Transmission using a JPEG2000-Like Architecture," IEEE Int. Conference on Image Processing, ICIP 2006, Atlanta, Georgia, October 8–11, 2006.

[3]   O. Bursalioglu, M. Fresia, G. Caire and V. Poor, "Lossy Joint Source-Channel Coding Using Raptor Codes," *International Journal of Digital Multimedia Broadcasting*, ID 124685, August 2008.

[4]   A. Abbasfar, D. Divsalar, and K. Yao, "Accumulate-Repeat-Accumulate Codes," *IEEE Transactions on Communications* **55** (April 2007): pp. 692–702.

[5]   A. Cohen, I. Daubechies, and J. Feaeveau, "Biorthogonal Bases of Compactly Supported Wavelets," *Pure Appl. Math.* **45** (1992): pp. 485–560.

[6]   D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practices*. Norwell, Massachusetts: Kluwer Academics Publishers, 2002.

[7]   S. McLaughlin, D. Neuhoff, and J. Ashley, "Optimal Binary Index Assignments for a Class of Equiprobable Scalar and Vector Quantizers," *IEEE Trans. Inform. Theory* **41** (November 1995): pp. 2031–2037.

[8]   N. Farvardin and V. Vaishampayan, "Optimal Quantizer Design for Noisy Channels: An Approach to Combined Source-Channel Coding," *IEEE Trans. Inform. Theory* **33** (November 1987): pp. 827–838.

H.   FIGURES
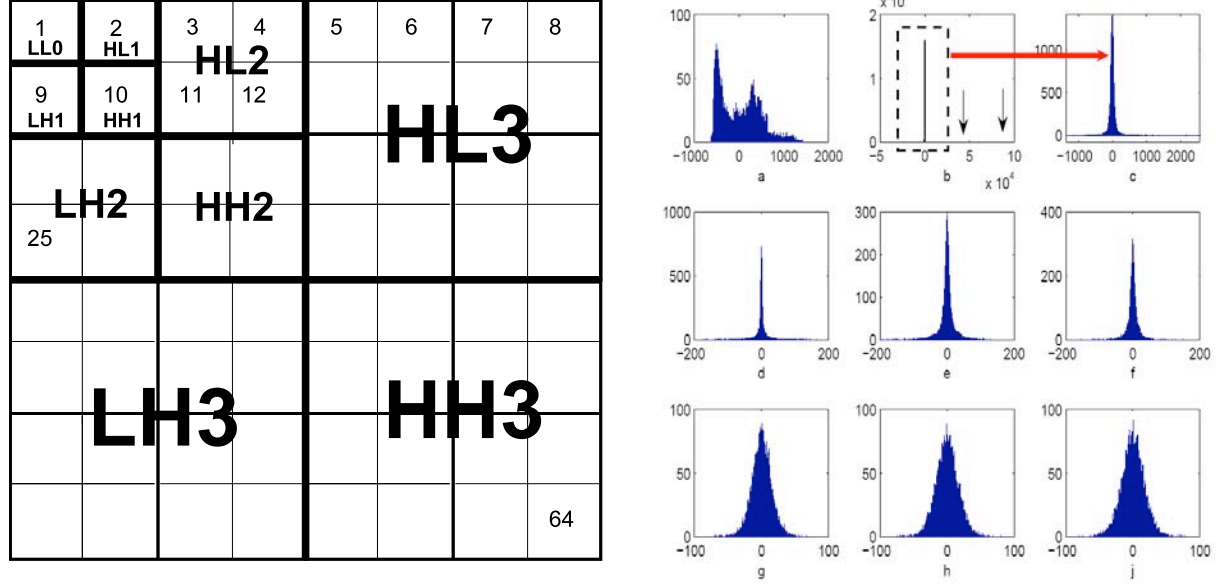
Figure 1. Subband decomposition of a DWT with W = 3 layers and DWT coefficient histograms. Figures a) b) c) refer to the LL0 subband without and with additional DCT.
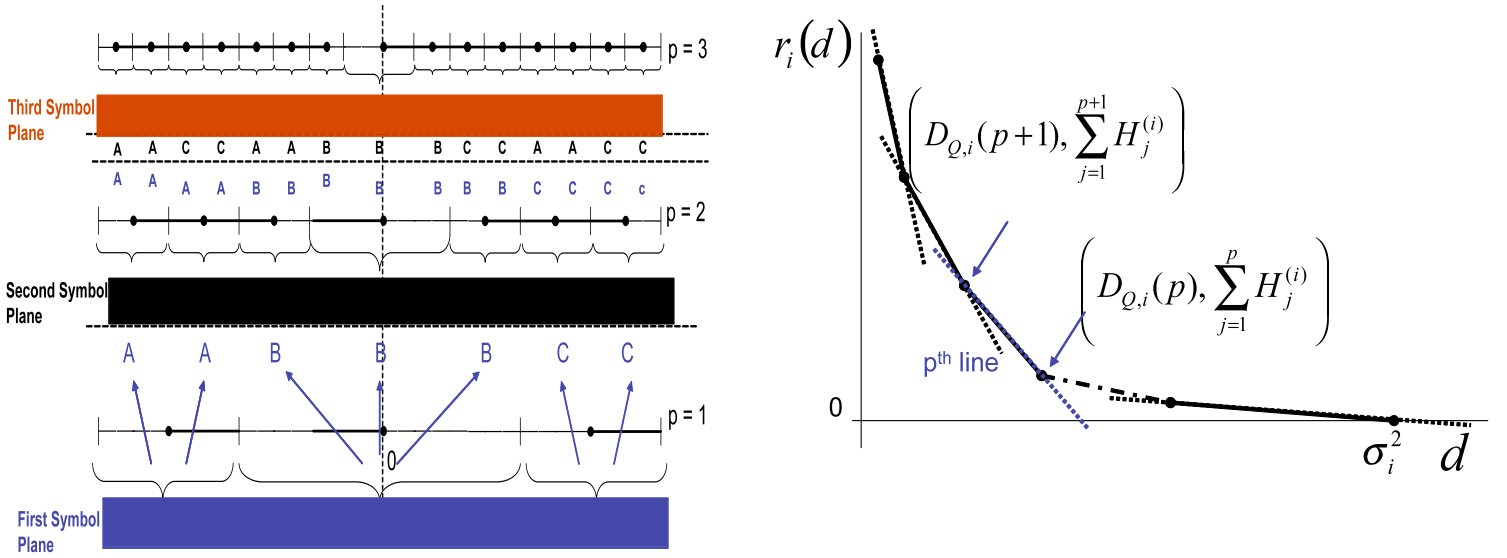
Figure 2. Embedded dead-zone scalar quantizer with ternary index labeling and corresponding piecewise linear convex operational rate-distortion function of the embedded quantizer.
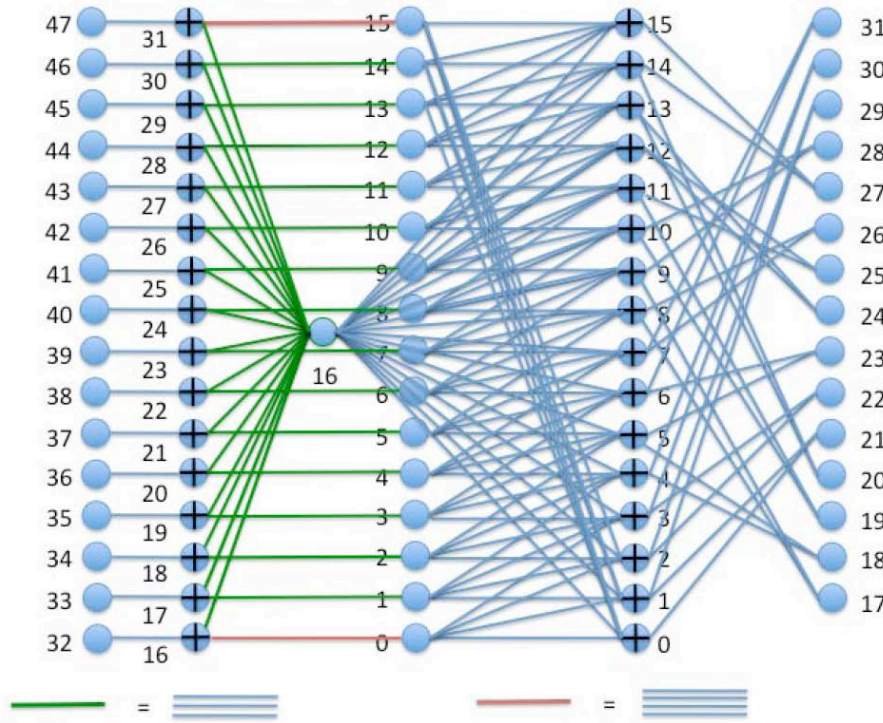
Figure 3. Protograph of the rate-compatible LDPC code family developed in this work.
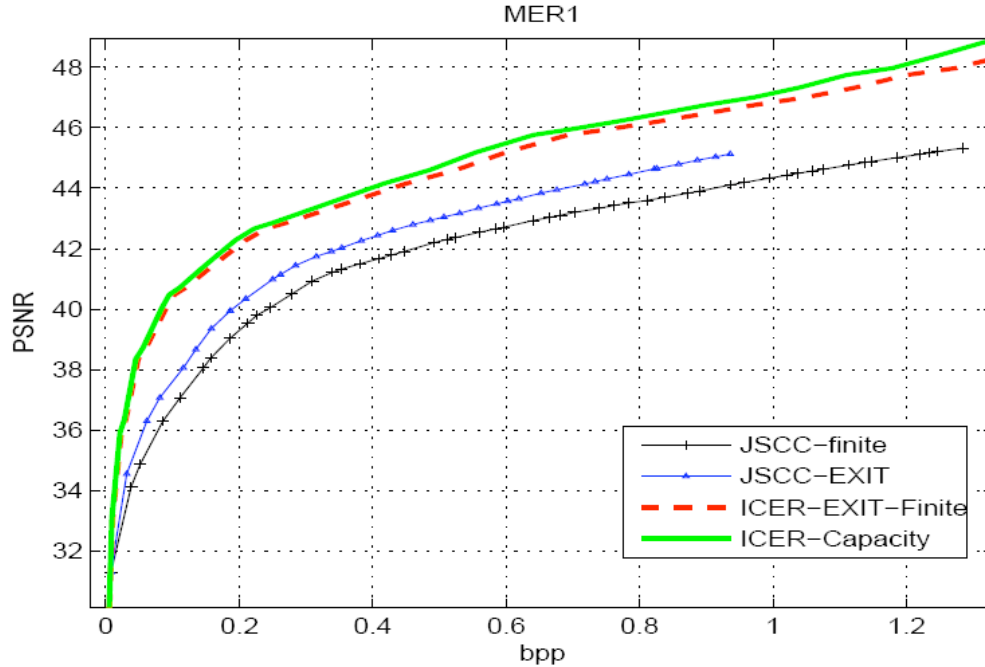
Figure 4. MSE distortion of the MER1 image over BI-AWGN channel at *Es/N0 = 3* dB, assuming infinite-length coding (convergence results provided by EXIT chart).

EXOPLANETS

SELF-HEALING, SELF-ADAPTIVE, LOW-POWER, HIGH-RESOLUTION ANALOG-TO-DIGITAL CONVERTERS FOR SPACE APPLICATIONS

Mohammad M. Mojarradi (PI), Instrument Electronics and Sensors Section (389)
Nazeeh I. Aranki, Flight System Avionics Section (345)
Bahar Jalali-Farahani, Department of Electrical Engineering, Arizona State University
Hugh Barnaby, Department of Electrical Engineering, Arizona State University
Gennady Gildenblat, Department of Electrical Engineering, Arizona State University

A.   OBJECTIVES

The objective of our work was to design an innovative, instrumentation-quality, 24-bit radiation-tolerant, self-healing and self-adaptive, analog-to-digital converter maintaining high resolution and low power consumption. The analog circuit was designed to work without degradation in performance over a wide temperature range of –230 to +120 C, so as to work reliably in the extreme lunar environment. In order to achieve this high resolution over the extended temperature range, we used several novel techniques such as CLS (correlated-level shifting) to maximize the signal power and reduce the effect of finite gain of Opamp, coarse and fine sampling to reduce the power consumption of the driver, chopper-stabilization to reduce the flicker noise, and Opamp offset. These techniques were embedded in the design of the front-end switched capacitor integrator. Moreover, a power-noise optimization method was employed in the architecture and system-level design of the sigma delta modulator. Healing and adaptivity features of the converter were embedded in the biasing circuit as well as the global system. A novel biasing circuit that combines the constant-gm method and the slew-rate enhancement technique were used to guarantee the minimum variation of both small signal and large signal parameters of Opamp. A global online calibration technique was developed that is capable of identifying the transfer function of the sigma delta modulator and hence compensate for the linear gain error. To ensure a successful design and simulation of the novel circuit, this project was empowered by the accurate compact models for CMOS. Novel CMOS compact models for low temperatures are being developed as part of the current project by the PSP model group at ASU.

B.   APPROACH AND RESULTS

The main objectives of this task are:

1.   Develop cryo compact models for bulk CMOS.
2.   Design a high-resolution (24 bits) analog-to-digital converter that operates under extreme temperature range of –230 to +120 C.

The approach taken for each task and the milestones are described here.

1.	Develop cryo compact models for bulk CMOS: The approach taken in the present study is to develop the compact model of MOS transistors for the extended temperature range, starting with the core model of the long-channel devices followed by the inclusion of the small-geometry effects as provided in the industry standard PSP model. In this way, the circuit designers can use the familiar tools with the only difference that the new transistor models are now validated for the extended temperature range.

To accomplish this objective, the core model was developed by including the incomplete impurity ionization in the surface-potential equation [1] followed by the development of the symmetric linearization method (theoretical foundation of the PSP model) in a form compatible with the new surface-potential equation [2]. The resulting core model has been verified by comparison with the low-temperature version of the Pao-Sah model also developed in the course of this work. The results of this comparison (see Figure 1) indicate a good agreement of the new core compact model and the Pao-Sah model thus validating the new approach.

2.	Design of the high-resolution analog-to-digital converter [3]: A second-order switched-capacitor sigma-delta modulator was designed and fabricated in IBM 130 nm technology. The sigma-delta technique was used since it provides the best trade-off between speed and resolution and can achieve high resolution by exploiting a very large oversampling ratio together with noise shaping technique. The designed ADC achieves 24 bits over 100 Hz signal bandwidth, which would scale down to 20 bits at 1 kHz. Several novel techniques were embedded in the design of the integrators to reduce the noise and ensure maximum signal-to-noise ratio. Figure 2 shows the switched capacitor implementation of the first integrator in the sigma delta modulator. This implementation has used the following novel technique:

a.	Coarse and fine sampling technique: This was done by the buffers and extra switches shown at the far left end of Figure 2. The buffers are not on the signal path and can be extremely low power. Applying this technique would relax the driving requirement of the source and reduce the power consumption of the driver significantly.

b.	Correlated-level shifting (CLS) was used in this design to enable a rail-to-rail operation of the switched capacitor integrator. The signal can therefore exploit the whole available range, which would maximize the signal-to-noise ratio. CLS also helps with reducing the error caused by the finite gain of Opamp and enables us to use less gain and hence less power in Opamp.

c.	Switched-capacitor integrator shown in Figure 2 also incorporated chopper stabilization technique to reduce the flicker noise and offset of the Opamp.

In order to maintain the performance of the ADC across a wide range of temperature, a wide-temperature biasing circuit was designed. The proposed biasing circuit combines a constant-gm biasing scheme with a slew-rate enhancement circuitry. The constant-gm circuit maintains the minimum changes in the transconductance and hence unity-gain-

frequency of the Opamp. The proposed slew-rate enhancement (SRE) technique, which is shown in Figure 3, maintains a constant slew-rate by injecting extra current when slewing condition is detected. Both small signal as well as large signal parameters are therefore kept constant over a wide range of temperature by combining the constant-gm and SRE techniques. Figure 4 shows the details of the SRE technique. Schematics of the Opamp and the pre-charge amplifier are shown in Figures 5 and 6 respectively.

C.   SIGNIFICANCE OF RESULTS

This research addresses the JPL SURP strategic challenge of "enabling robotic missions to scientifically interesting extreme environments" by providing novel approaches for high-performance data acquisition system working in these extreme conditions. Temperature- and radiation-tolerant as well as survivability are key features required for electronic circuits used in NASA missions. This research was a collaborative effort between modeling and circuit design groups at ASU. Accurate cryo PSP models of bulk CMOS were developed through the course of this project, which were then used to verify the performance of the circuits across the wide range of temperature. Incorporating several novel techniques, a high-resolution sigma-delta ADC was designed that achieves 24 bits at 100 Hz with scalable resolution of 20 bits at 1 kHz. Through using a novel biasing scheme that combines the constant-gm with a slew-rate-enhancement technique, minimal variation in small as well as large signal parameters of the Opamp was ensured. Hence, performance of the ADC is maintained across the wide range of temperature. In addition to the development of the wide-temperature ADC, the approach to compact modeling at extreme temperatures as well as the enablement of radiation effects modeling for circuit design can be extremely advantageous to JPL. Through the use of these models, IC developers at JPL may be able to design for environments characterized both by extreme temperatures and high level of radiation.

D.   FINANCIAL STATUS

The total funding for this task was $100,000, all of which has been expended.

E.   PUBLICATIONS AND PRESENTATIONS

[1]   G. Gildenblat, Z. Zhu, and C. C. McAndrew, "Surface Potential Equation for Bulk MOSFET," *Solid-State Electronics* **53**, pp. 11–13, 2009.

[2]   Z. Zhu and G. Gildenblat, "Symmetrically Linearized Charge-Sheet Model for the Extended Temperature Range," *Solid-State Electronics*, in press.

[3]   S. Gopal Krishna and B. Jalali-Farahani, "Fast Settling Slew Rate Enhancement Technique for Operational Amplifiers," submitted to IEEE International Midwest Symposium on Circuits and Systems, 2010.

F.   REFERENCES

[1]   G. Gildenblat, Z. Zhu, and C. C. McAndrew, "Surface Potential Equation for Bulk MOSFET," *Solid-State Electronics* **53**, pp. 11–13, 2009.

[2]   Z. Zhu and G. Gildenblat, "Symmetrically Linearized Charge-Sheet Model for the Extended Temperature Range," *Solid-State Electronics*, in press.

[3]   S. G. Krishna, "High-Resolution Sigma-Delta Modulator for Spectroscopy," M.Sc thesis, December 2009.

G.   FIGURES
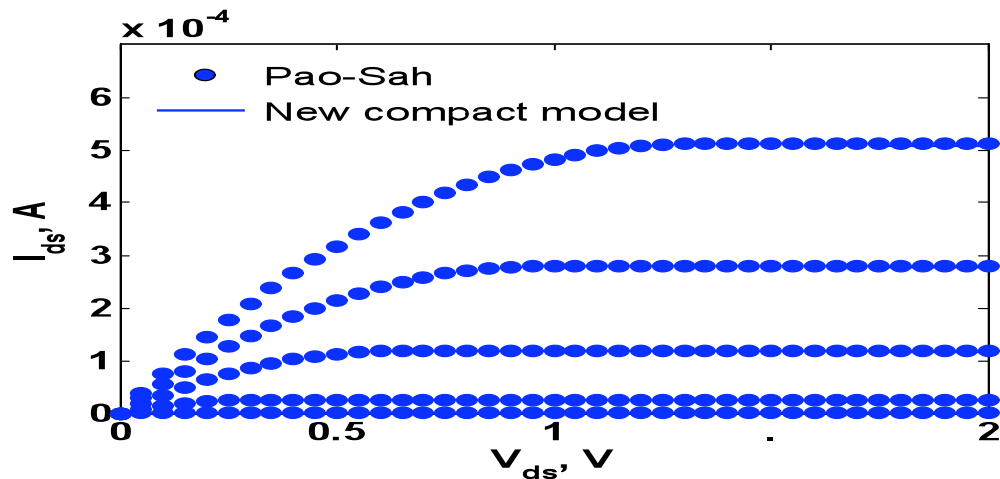
Figure 1. Agreement of Pao-Sah model with the new compact model.
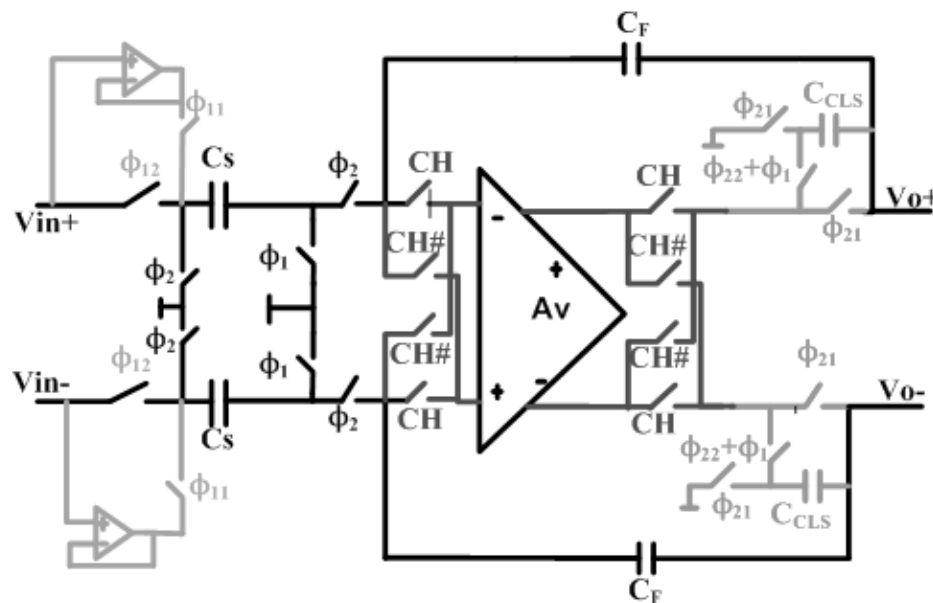
Figure 2. Switched-capacitor implementation of the first integrator in the sigma-delta ADC.
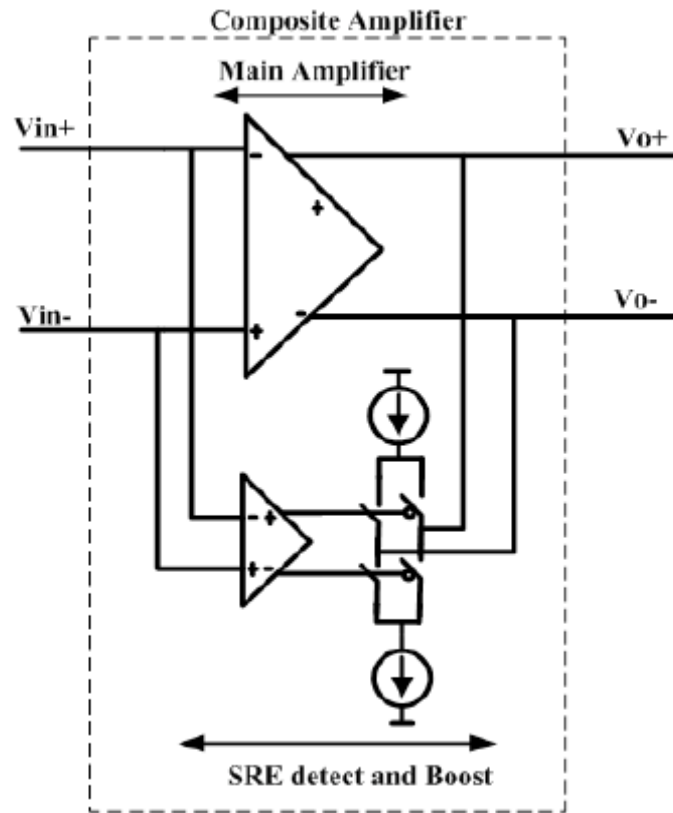
Figure 3. Opamp with the slew-rate-enhancement (SRE) circuitry.
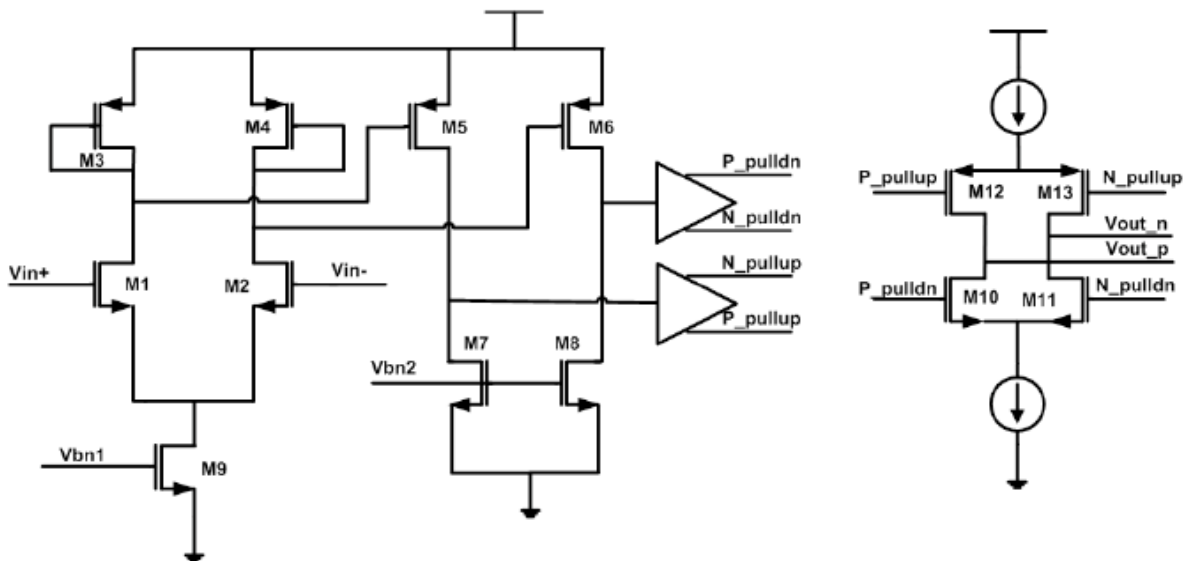
Figure 4. Schematic of the auxiliary Opamp shown in Figure 3 used for detecting the slew rate condition and injecting the extra current.
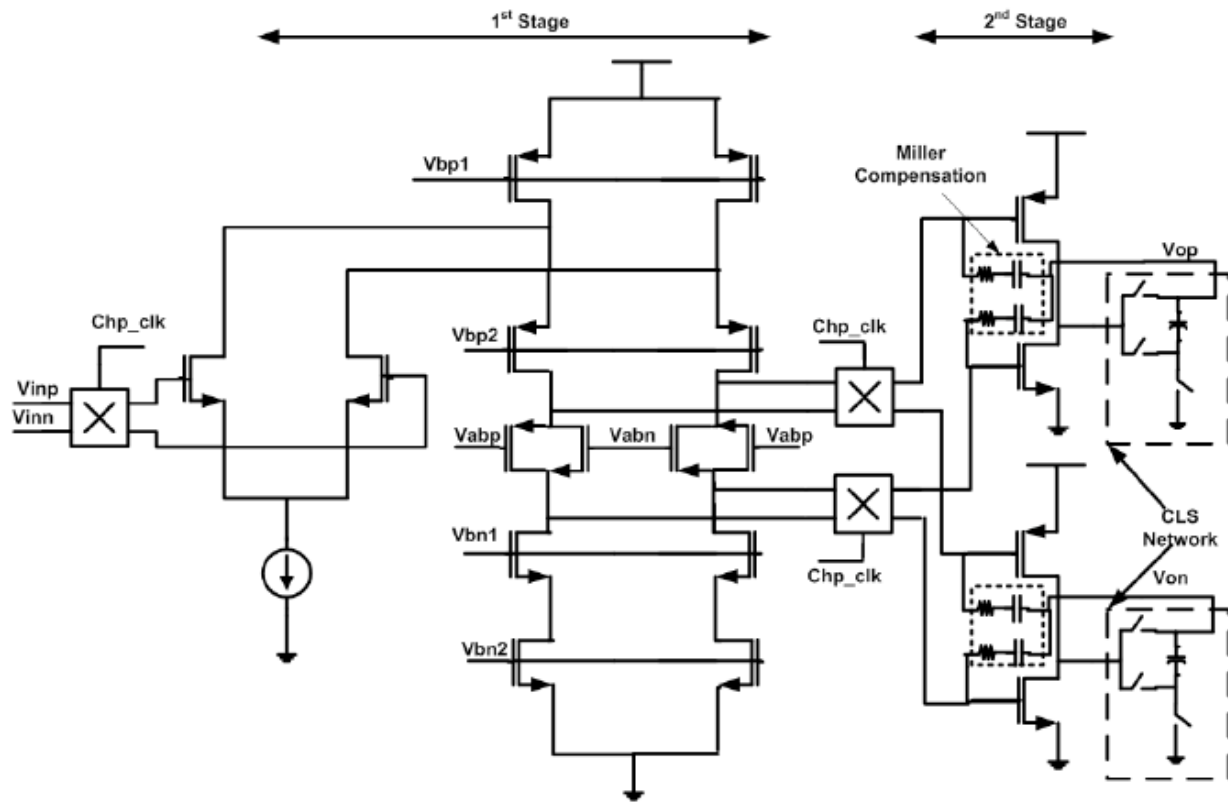
Figure 5. Schematic of the Opamp with the chopping and correlated-level-shifting switches.
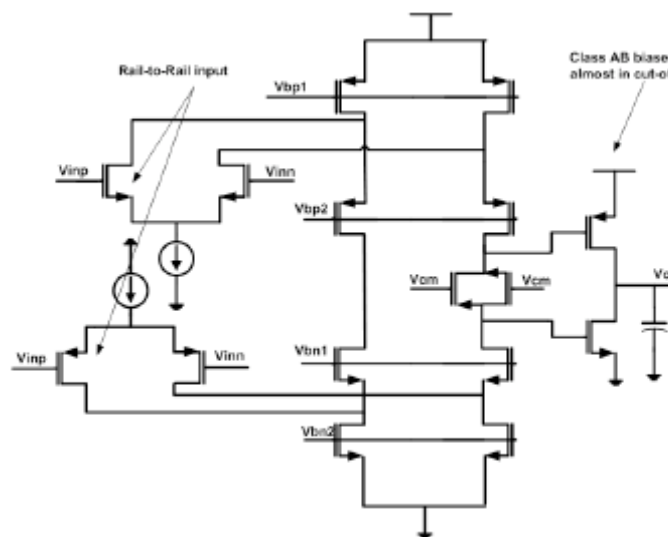
Figure 6. Schematic of the precharge amplifiers.

DRILLING TECHNOLOGY FOR A MARS SAMPLE RETURN:
A THEORETICAL INVESTIGATION

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1327

Samad A. Hayati (PI), Mars Exploration Directorate (600)
Lori R. Shiraishi, Instrument Mechanical Engineering Section (355)
Timothy J. Szwarc, Department of Aeronautics and Astronautics, Stanford University
Scott Hubbard, Department of Aeronautics and Astronautics, Stanford University
Brian J. Cantwell, Department of Aeronautics and Astronautics, Stanford University
Richard M. Christensen, Department of Aeronautics and Astronautics, Stanford University

A.   OBJECTIVES

The processes of coring and drilling are central to energy production, climate change research, and space exploration. While new materials and equipment have allowed drilling depths, efficiencies, and penetration rates to increase, the understanding of coring and drilling mechanics has not matured in step. In order to produce coring and drilling systems that can be optimized for power efficiency or other specifications, it is essential to determine the relations between inputs, such as hardware design and operation, and outputs, such as the drill's rate of penetration and power efficiency. Rather than advancing the technical readiness level of a particular device, this research was aimed at providing a general knowledge base from which improvements to coring and drilling systems can be made.

The objective of this task was to gain an understanding of the current state of knowledge in the mechanics of coring and drilling, as well as the scientific and engineering requirements of a Mars sample return (MSR) campaign. In doing so, the key inputs and outputs of an MSR coring system were identified and quantified. To prepare for the next stages of work, which will focus mainly on experimentation, a testbed was constructed that allows for quantitative links between inputs and outputs to be measured. The results from these tests, along with knowledge gained through the survey of materials and structural literature, will be used to create models between inputs and outputs of coring and drilling performance. As the models are created from experimental data, a theoretical analysis of rock mechanics will be used to explain the relationships between inputs and outputs. Finally, these models can be used to optimize hardware design and implementation to achieve goals specified by mission requirements.

Throughout the research, care was taken to identify novel technologies and methods of coring or drilling with advantages over traditional designs for possible future development and use.

B.   APPROACH AND RESULTS

1.   Survey of Spacecraft Sample Acquisition and Mission Literature

An extensive literature survey was conducted to ascertain the current state of coring and drilling technologies, including the strengths and weaknesses of existing tools; and the engineering and science requirements likely to be placed upon an MSR coring system.

a.   *Current state of drilling technology.* Among the many acquisition systems investigated, particular attention was paid to studies of the Mars Science Laboratory (MSL) powder-acquiring drill and the ExoMars corer designed by Agenzia Spaziale Italiana (ASI). The former, which is designed to acquire powdered rock from exposed rocks on Mars, is subject to many requirements similar to those of a potential future MSR corer. Due to the time constraints of the MSL project, the powder-acquiring drill was designed and improved as parts and techniques were tested and iterated. No theoretical modeling was performed.

During this investigation, meetings were conducted at JPL with experts in instrument design to exchange ideas regarding the challenges of core sample acquisition and handling. Among the key lessons learned during the initial investigation were the specific roles that rotation and percussion play in penetration. While rotation is sufficient for penetration in many materials, such as wood or glass, it is percussion that drives penetration in rock. Percussion creates stress waves near the interface of the bit and rock, leading to fracture of the nearby rock. Rotation helps to clear cut material from this interface and provide the bit with fresh material to impact.

In addition, study of the ASI corer provided lessons from an existing coring device, although the device is not capable of penetrating rock as strong as some rocks to be sampled by the proposed MSR. Therefore, many significant and fundamental changes would be required to adapt the ASI corer to JPL's requirements for an MSR. Bit design and methods for extracting cores from boreholes were among the lessons learned through study of the ExoMars corer. Finally, other previously built coring devices, including the Honeybee Corer/Abrader Tool and Minicorer, were analyzed to reveal the difficulties in producing hardware for core sample acquisition.

b.   *Proposed science and engineering requirements of an MSR mission.* To understand the needs and limitations of possible future sample acquisition systems, preliminary MSR mission requirements were studied. At present, the expected requirements for an MSR sample acquisition system include collection of about 35 samples, approximately 1 cm in diameter and 5 cm in length, from rocks ranging from hardness level 2 (phyllosilicates) to 7 on the Mohs scale [1]. Because of the Martian environment, all coring must be done without the use of any lubricants, and mass and power consumption must be minimized as with any mission. Finally, because of the limited mass of a rover, the axial weight on bit (WOB) could not be arbitrarily large and would be limited to a percentage of the spacecraft's weight on Mars and also would depend on the angle the bit makes to the vertical during operation.

In addition, the expectations of the science community were studied to determine the characteristics of a sample that would be most important in providing valuable scientific material for analysis. Because sedimentary rocks are of great interest due to the clues they reveal about the history of water on Mars, the preservation of the layered content of a sample would be of chief importance. Therefore, acquiring samples in such a way that the structural integrity of the core would suffer minimal damage is essential. Powdered rock or a collection of small rock fragments would not provide the same level of detail as would a

solid, contiguous core. Since the layers within a sample may have vastly different structural properties, including a large range of possible Mohs hardness levels, it is also important for a coring system to be able to adjust its operating parameters during penetration. For example, being able to adjust rotation speed based on feedback, such as power consumption, in order to penetrate forcefully or gently where necessary, would help to prevent cores from being destroyed. Finally, to preserve a core sample's chemistry and to prevent the evolution of volatile substances, temperature control during penetration would be essential; more specifically, preventing a sample's temperature from rising above –20 °C if possible [2].

2. Computer Simulation of Rock and Stress Deformation

A computer simulation was performed to analyze the stresses and displacements of an arbitrary bit geometry on the face of a rock. This simulation is based primarily on classical solutions to load cases. At present, this simulation can only accommodate static loading, but further work may be aimed at allowing dynamic loads that cause stress waves to propagate through the material. Additional work may also allow for arbitrary rock geometries to be loaded, rather than a semi-infinite medium. This simulation was used to model typical drill bit tooth geometries and identify the locations of highest stress in the rock, including tensile loads near the rim of the indentation that are likely to cause failure of brittle materials.

3. Testbed Construction

A testbed was constructed, which was based on a successful testbed at JPL's planetary sample acquisition lab. This setup allows for a commercially available hammer drill to be loaded with any amount of axial loading and to translate smoothly along a linear track. Currently, capabilities exist to measure rotation rate, percussion rate, rate of penetration, and total power consumption, with new sensors to be installed in the future. An experiential test matrix was constructed for the testbed in order to provide adequate data points to determine the relations between the inputs and outputs of the drill system. The first set of experiments in this plan was conducted in order to determine the link between bit diameter and rate of penetration (ROP). WOB was scaled with the cross-sectional area of the bit in order to maintain a constant pressure on the bit face. Power was maintained at 100% of the drill's capability. Note that this does not necessarily mean that a constant net power was applied to the bit, so additional sensors will need to be fitted to the testbed before this quantity can be determined and controlled. As can be seen in Figure 1, the results of this test show that ROP appears to be constant with bit diameter, given that all other quantities were held equal. Similar tests will be completed with different pressures on bit and different rocks to test the validity of this result. In addition, once the proper sensors are fitted, percussive power will be held constant through a test, since it is speculated that percussive power is the input most directly linked to rock failure and penetration rate. Future testing will include measuring ROP against changes in WOB for a given bit size, in addition to ROP against power consumption in order to identify more trends.

To complement the testbed, a computer program was written to determine the percussive frequency of a drill test by analyzing a sound recording of the test. The results of this program were compared to a tachometer to verify the accuracy of the program. The advantage of the program over the use of a tachometer is that the program continuously measures and records frequencies, which can be saved and used later to calculate quantities such as power consumption. Figure 2 shows the results of a typical drill test.

C. SIGNIFICANCE OF RESULTS

Because the majority of the work completed thus far was an investigation of the challenges of coring and drilling, the most significant results of the study are those that will guide the direction of future experimental work in preparation for the proposed MSR. The direction of investigation will help to produce test data that reveal the most about performance characteristics specified in the proposed mission requirements.

As stated earlier, rotation and percussion each serve important, yet distinct roles in the penetration of rock. Therefore, the effects of rotation and percussion must be studied individually. Future work will be aimed at measuring power inputs to the rotation and percussion mechanisms of the testbed in order to determine trends in output performance based on these powers. Since the effects that rotation and percussion have on the rock are coupled, it is not possible to isolate their effects simply by using one form of motion at a time. Therefore, a data acquisition system is being designed and fitted to the testbed to simultaneously measure and control rotary and percussive power inputs to the system while performing a typical drilling or coring test. To simulate the energy dissipation characteristics of a drill or corer operating on Mars, the testbed uses rocks that are mounted in a bed of sand and gravel rather than rigidly fixed to a surface.

Also identified was the cause of the tendency of penetration to slow with increasing depth. The cause of this trend was not the wear of the cutting edge of the bit, but primarily the quantity of powdered material that builds up during penetration under the bit, preventing the bit from impacting new rock. Typically, increasing the WOB can help to overcome the decrease in penetration rate [3], but mission requirements preclude this solution past a certain threshold. Further work will be done to analyze the role of rotation in clearing material from under the bit in order to model and eventually minimize this effect. Additionally, vibration aids in the clearing of cut material, and modeling particle flow may also help to find ways to minimize reductions in penetration rate.

Previous investigation was crucial in creating an experimental test plan to look at the inputs and outputs most relevant to the projected science and mission requirements of MSR. Early tests will seek to identify relationships between energy consumed to drill a hole and the volume of material extracted. Should consumption not scale linearly with extracted volume, for example, then gains in efficiency can be realized by drilling pilot holes or perforation patterns prior to the final penetration for sample acquisition. In addition, WOB and the frequencies of rotation and percussion will be altered to test not only energy consumption, but also the effects on the structural integrity of delicate or layered rock formations. Finally, heat conduction will be studied, either with optical sensors to measure surface temperature of the bit and rock, or with implanted sensors to model and help minimize thermal alterations to the sample. The experimental plan will be adapted as required based on trends noticed during ongoing analysis of the data.

The first set of experimental results has shown that if the pressure on the bit is held constant, ROP is constant despite varying bit diameter. This result agrees with the notion that penetration is achieved primarily by percussive impact to the rock. Since the pressure on the bit remains constant, so too does the force beneath a unit area of bit. A more in-depth analysis based on net power imparted to the bit will be needed to say more about this relationship.

Already, the results of this investigation are of use to JPL in preliminary planning for an MSR. Aside from new technologies that hold promise for mission design, to be discussed later in this report, the need for sampling algorithms tailored to specific rocks has been made obvious by this research. As experimentation and modeling continue, input parameters will be optimized according to hardness, friability, or other characteristics of the material to be sampled in order to prevent damage to the sample without using unnecessary amounts of energy or time in obtaining the sample.

## D.	NEW TECHNOLOGY

The concept of a laser-perforated coring system is currently under the patenting process and may be developed further in the future. The system couples a traditional coring device with a pulsed laser capable of ablating rock. The pulsed laser is used to produce a pattern of small holes in order to weaken selected areas of the rock. It is hypothesized that the energy savings gained by penetrating this weakened rock with a coring or drill bit would be greater than the energy cost of operating the laser, thus leading to a net savings in energy consumption.

An additional advantage of the laser-perforated system would be the ability to guide the location of rock failure, which would be used to preserve the structural integrity of core samples. Because the rock surrounding a core sample would be significantly weakened by the laser, the structural integrity of the sample would be relatively higher and would be left intact during extraction. Guiding the failure of the rock in this way would not only produce stronger cores, but would also produce cores that are more perfectly cylindrical and therefore easier for the sample acquisition system to handle and transport.

Finally, since the rate of penetration would be higher in the weakened rock than it would otherwise be for a given WOB and rotation speed, heat generation would be reduced. This is because heat is generated by friction as the bit is rotated within the rock, and reducing drilling time would reduce energy losses to friction. Heat conduction would be reduced because heated rock is cut and removed from the rock more quickly than the original, unperforated rock would otherwise be. It is hypothesized that heat added by the laser would be localized to small regions where the laser beam strikes the rock, and these regions would be quickly removed by the bit before being able to conduct heat and alter the contents of the sample.

## E.	FINANCIAL STATUS

The total funding for this task was $195,700, of which $194,591 has been expended.

## F.	ACKNOWLEDGEMENTS

G.   PUBLICATIONS AND PRESENTATIONS

[A]   Timothy Szwarc, Scott Hubbard, Brian Cantwell, and Richard Christensen, "Core Acquisition and Handling for Mars Sample Return," presentation to invited guests, sponsored by JPL, August 24, 2009.

[B]   Timothy Szwarc, Scott Hubbard, Brian Cantwell, and Richard Christensen, "Principles and Future Directions of Mars Sample Acquisition," presentation to invited guests, sponsored by JPL, February 16, 2010.

[C]   Timothy Szwarc, Scott Hubbard, Brian Cantwell, and Richard Christensen, "Modeling of Coring Mechanics for Optimization of Mars Sample Acquisition Hardware," poster session sponsored by Stanford University, April 26, 2010.

H.   REFERENCES

[1]   IMARS Working Group, Preliminary Planning for an International Mars Sample Return Mission, June 1, 2008.

[2]   Lars Borg, David Des Marais, et al., "Science Priorities for Mars Sample Return," Mars Exploration Program Next Decade Science Analysis Group, March 15, 2008.

[3]   Kris A. Zacny and George A. Cooper, "Coring Basalt under Mars Low Pressure Conditions," *The International Journal of Mars Science and Exploration* **3** (March 9, 2007): pp. 1–11.

I.   FIGURES
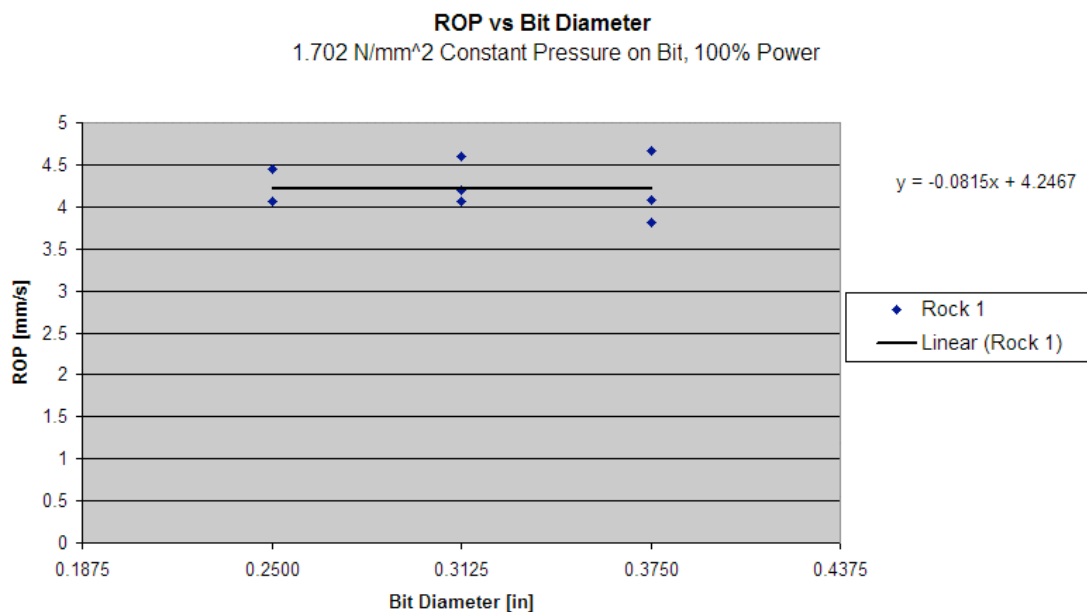
Figure 1. Rate of penetration versus bit diameter.
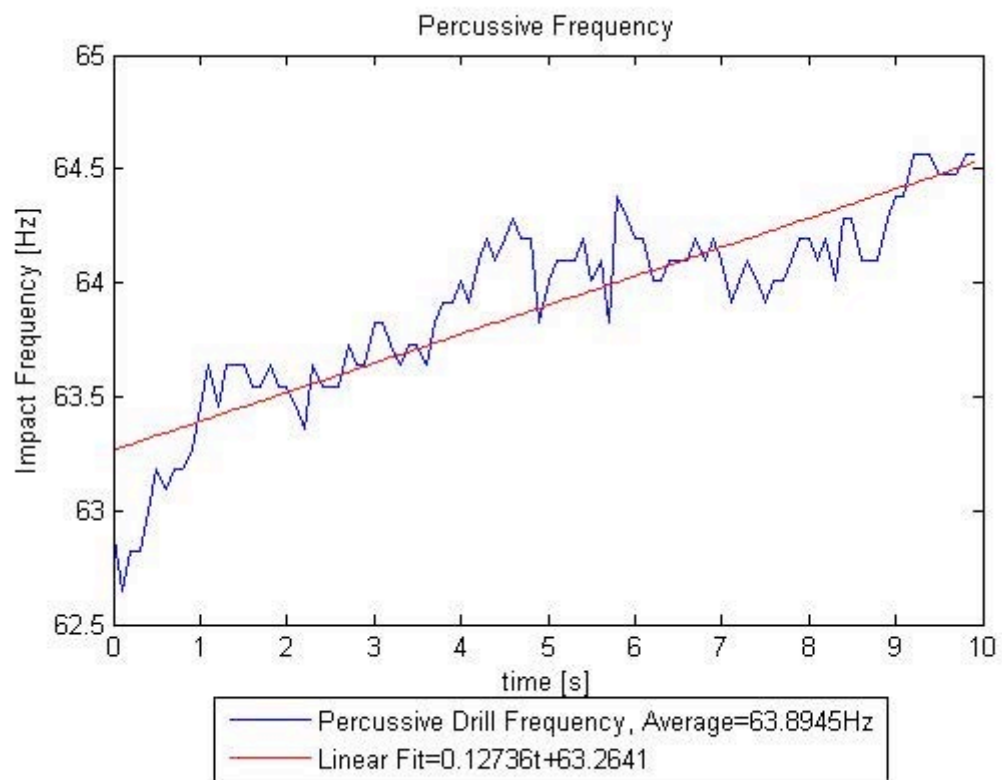
Figure 2. Percussive frequency versus duration of test.

GEOMETRY AND STRUCTURE
OF OUR UNIVERSE

COMPLIANT TASK EXECUTION AND LEARNING FOR SAFE-MIXED INITIATIVE HUMAN-ROBOT OPERATIONS

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1309

David S. Mittman (PI), Planning and Execution Systems Section (317)
Michel D. Ingham, Flight Software and Data Systems (316)
Vandana Verma, Robotic Software Systems Section (347)
Brian C. Williams (Co-PI), Department of Aeronautics and Astronautics and Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology
Patrick R. Conrad, Department of Aeronautics and Astronautics and Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology
Shuonan Dong, Department of Aeronautics and Astronautics and Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology
Julie A. Shah, Department of Aeronautics and Astronautics and Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology

A. OBJECTIVES

The objective was to deliver a novel task execution capability that enhances the ability of in situ crewmembers to function independently from Earth by enabling safe and efficient interaction with automated systems. This task execution capability provides the ability to (1) map goal-directed commands from humans into safe, compliant, automated actions, (2) quickly and safely respond to human commands and actions during task execution, and (3) specify complex motions through teaching by demonstration. Our task execution capability is intended to complement current work in architectures for goal-directed commanding [1]. Results are applicable to future surface robotic systems, and we have demonstrated capabilities on JPL's All-Terrain Hex-Limbed Extra-Terrestrial Explorer (ATHLETE) robot.

B. APPROACH AND RESULTS

The goal of this task was to provide a more natural, common-sense way of interacting with the next generation of NASA robotic systems. One example of such a robot is ATHLETE, a vehicle developed at JPL to support robotic and human missions on the surface of the Moon. Previously, robots have been simple enough to be controlled through teleoperation or through scripts composed of many low-level motion commands. However, more complex robots, like ATHLETE (Figure 1) or the Vecna BEAR battlefield robot, have too many degrees of freedom for direct teleoperation or commanding to be practical. For example, a script for the ATHLETE rover, a 36-degree-of-freedom robot, are composed of many commands such as "move joints speed 0.8 absolute motion limb1 hipyaw 0.510 hippitch 0.221…"

To enable more efficient interaction, we have developed a different control paradigm that incorporates verbal commands, shared written instructions, and demonstration by example. These modes of communication occur naturally to humans during collaborative tasks. This control paradigm allows the user to specify intuitive, goal-directed commands, and then teach the robot what these commands mean through demonstrations. An intelligent executive then interprets these commands in the context of shared written instructions, and transforms them into low-level motion commands to the robot. There were two key contributions on this project: mapping goal-directed commands into safe, compliant, automated actions, and specifying complex spatial motions through teleoperated demonstration.

1.   Mapping goal-directed commands into safe, compliant, automated actions

Allowing the user to specify intuitive, goal-directed commands to the robot requires an intelligent executive to interpret the commands and transform them into low-level machine instructions. The process begins with the user giving directions to the robot in the Reactive Model-based Programming Language (RMPL) developed by the MIT's Model-based Embedded Robotic Systems (MERS) group [2]. This language allows high-level instructions at a level of abstraction that is more natural for the user, such as "move the box," in a convenient Java-like syntax. The user can also provide constraints that the system needs to obey without giving specific instructions on how to do so. We mostly deal with temporal constraints: restrictions on the duration of activities and completion deadlines. The executive is then responsible for reasoning through the constraint set forward in the plan and executing the activities correctly.

The executive performs constraint reasoning to determine all the implications of the user's requests and then adjusts the execution in real-time in response to the outcomes of events, making it more robust to disturbances. The executive is given explicit direction about what the user requires, allowing it to determine the total set of possible plans and schedules that meet those requirements. Then, the executive is empowered to shift between different executions rather than being bound to a single type of reasoning or a single plan provided before even beginning. A reactive executive, which can adjust the execution sequence on the fly as plan unfolds, can guarantee the correct execution of a plan even with some disturbances. To make this possible, the executive is broken into a compiler and a dispatcher. The input plan in RMPL is converted into a temporal constraint formalism to allow reasoning on it, and the compiler performs a type of reasoning on it, reformulating it into a dispatchable form. The dispatcher works with the dispatchable form to efficiently schedule the start and end of the activities in the task, adjusting the plan as it observes the start and end of events.

In this work, we tested two related executives on ATHLETE, named Chaski [3] and Drake [4]. Both allow temporal reasoning with choices that the robot is allowed make: Chaski focuses on multi-agent collaboration and Drake focuses on allowing arbitrary choices between subcomponents of a plan. Chaski was successfully demonstrated to perform tight temporal synchronization of two of ATHLETE's limbs, modeled as independent robots, even with uncertainty in how long various actions would take. Specifically, the two limbs had to be synchronized to pass a rock held by one limb into a sample container held by the second, which requires that both limbs act at the same time.

The demonstration of Drake involved a plan with an optional high-value task that could be inserted into the plan, but only if sufficient time remained after an uncertain duration. The executive could not predict in advance if there would be sufficient time for the extra task, requiring that it wait and see and adjust afterward. The plan used for our demonstrations is shown in Figure 2 along with the plan described in Reactive Model-based Programming Language. We were able to successfully demonstrate Drake's capability of online decision-making.

2.   Specifying complex spatial motions through teleoperated demonstration

Sometimes specifying low-level commands to achieve complex motions, such as "move the box," can be a challenge. We reduce time and effort required to do this by teaching actions through demonstration. For ATHLETE, we use an interface device called Tele-robotic ATHLETE Controller for Kinematics (TRACK, aka ADMIN) [5], developed by the Distributed Robotics Lab at MIT, to teleoperate the robot through each desired activity several times.

Based on the demonstrations, our learning algorithm generalizes from the examples a set of motions that will likely achieve the action of "move the box." This is compactly encoded in a representation we call a probabilistic flow tube, as illustrated in Figure 3. Flow tubes have been used in the context of planning with continuous motions to represent sets of trajectories with common characteristics [6,7]. We have incorporated this concept to motion learning by introducing the probabilistic version. The learned probabilistic flow tube represents flexibility in the motions the robot can choose to perform, enabling it, for example, to recover from disturbances. The width of the flow tube indicates areas with more or less flexibility.

A key feature of our learning system is the ability to autonomously determine what parameters, if any, are relevant in a demonstrated motion. We use the term parameter to describe qualitative features of a motion. For example, a relative motion to move an object two feet to the right has two relevant parameters: the initial location of the object, and that the relative motion distance is 2 feet. Parameter identification uses clustering to determine if patterns exist across different training samples at object contact points. Our current system considers the following types of motions: (1) absolute motions, such as "move to the hole," (2) motions relative to an object, such as "move to the box," (3) motions relative to the robot, such as "move two feet to the right," (4) or unparameterized motions, such as "wave arm." Without any prior knowledge about what type of motion a demonstration might be, the system is able to determine which relevant parameters, if any, exist in the motion by observing patterns across training demonstrations. We assume the locations of the objects and robot in the environment can be determined through sensing, such as GPS or vision processing.

During demonstrated training, the initial positions of objects in the environment are not fixed or predefined, but rather sensed when needed. When the robot is prompted to execute the learned motion in a new situation, any object positions identified as relevant parameters are sensed, and the system generates a probabilistic flow tube using training trajectories with initial conditions similar to those observed. This flow tube is scale- and rotation-normalized to be used in the new situation.

We first tested our system in a 2-dimensional environment, in which there were located some randomly positioned objects (a red box and a green bin) and an absolute location marked by x. The motions we used to test the system included "put object in bin," "move object left 1," and "move object to x." We trained the system by clicking and dragging objects around. Figure 4 (left) shows an example of a resulting learned motion generated under previously unseen initial conditions.

During our on-site demonstrations at the JPL Mars Yard, we were able to show the feasibility of our motion learning system on ATHLETE. Using the TRACK interface device, we taught the robot to perform the "move box on top of platform" motion five times, each time with the box and platform in different initial locations. We then moved the box and platform to new locations, and asked ATHLETE to autonomously perform the motion. Figure 4 (right) shows the trajectory of the robot limb end-effector during demonstrated motions and autonomous execution of the "move box on top of platform" motion.

C.   SIGNIFICANCE OF RESULTS

We demonstrated two task executives, Chaski and Drake, on the ATHLETE platform, as well as the ability to teach the robot complex actions. Currently, intelligent task executives are not widely adopted in space applications. While the tasks we performed during the tests were short and mostly serve as proofs of concepts, demonstrating on advanced hardware platforms like ATHLETE raises the technology readiness level of our capabilities, and will aid in making the technology feasible for space applications. Furthermore, it lends weight to previous simulated experiments indicating the usefulness and performance of the algorithms behind the two executives.

The ability for a user to teach a robot some desired motions allows for more intuitive interactions. Our demonstration of motion learning on ATHLETE shows the potential feasibility and benefits of this mode of operation in future human-robot missions. In the box-moving example, we showed that only a few demonstrations might be necessary to obtain desirable generalized motions that can be autonomously executed in a wide range of new situations.

D.   FINANCIAL STATUS

The total funding for this task was $100,000, of which $91,264 has been expended.

E.   ACKNOWLEDGEMENTS

F.   PUBLICATIONS

None.

G.   REFERENCES

[1]   Daniel D. Dvorak, Michel D. Ingham, J. Richard Morris, and John Gersh, "Goal-Based Operations: An Overview," *Proceedings of the Infotech@Aerospace Conference and Exhibit*, Rohnert Park, California, May 7–10, 2007.

[2]   Brian C. Williams, Michel D. Ingham, Seung H. Chung, and Paul H. Elliott, "Model-based Programming of Intelligent Embedded Systems and Robotic Space Explorers," Invited Paper, *Proceedings of the IEEE: Special Issue on Modeling and Design of Embedded Software* **91** (January 2003): pp. 212–237.

[3]   Julie A. Shah, Patrick R. Conrad, and Brian C. Williams, "Fast Distributed Multi-agent Plan Execution with Dynamic Task Assignment and Scheduling," International Conference on Automated Planning and Scheduling, Thessaloniki, Greece, September 2009.

[4]   Patrick R. Conrad, Julie A. Shah, and Brian C. Williams, "Flexible Execution of Plans with Choice," International Conference on Automated Planning and Scheduling, Thessaloniki, Greece, September 2009.

[5]   Marsette Vona, "ATHLETE Direct Manipulation Interface," September 13, 2008. http://www.mit.edu/~vona/ADMIN

[6]   Andreas G. Hofmann and Brian C. Williams, "Exploiting Spatial and Temporal Flexibility for Plan Execution of Hybrid, Under-Actuated Systems," *Proceedings of the 21st National Conference on Artificial Intelligence*, Boston, Massachusetts, July 2006, pp. 948-955.

[7]   Hui Li and Brian C. Williams, "Generative Planning for Hybrid Systems based on Flow Tubes," Proceedings of the International Conference on Automated Planning and Scheduling (ICAPS), Sydney, Australia, 2008.

H. FIGURES

Figure 1. ATHLETE rover.
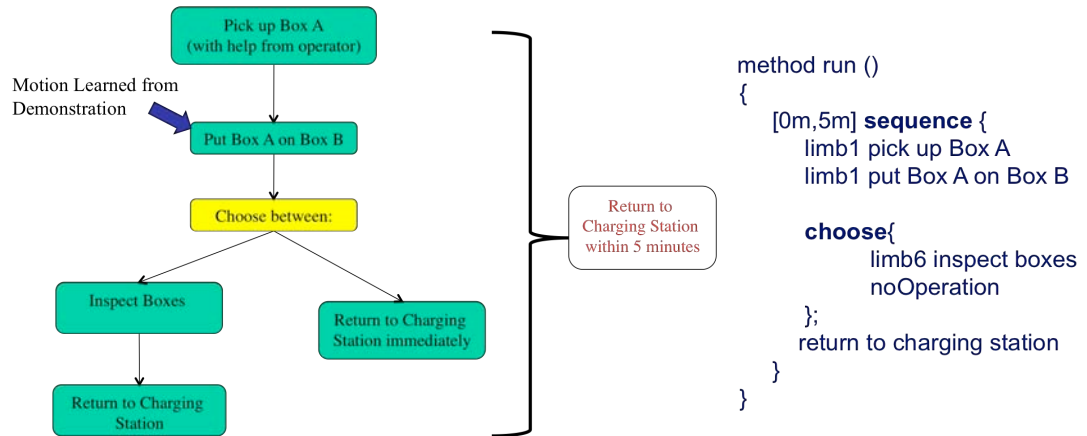
Figure 2. Drake demonstration plan and RMPL description.
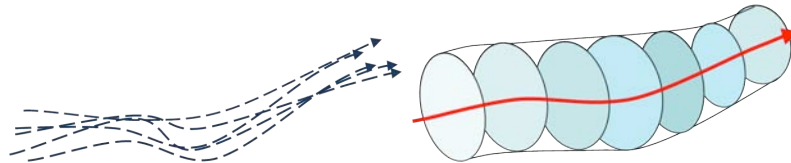
Figure 3. The probabilistic flow tube on the right is a learned generalization of the demonstrated motions on the left.
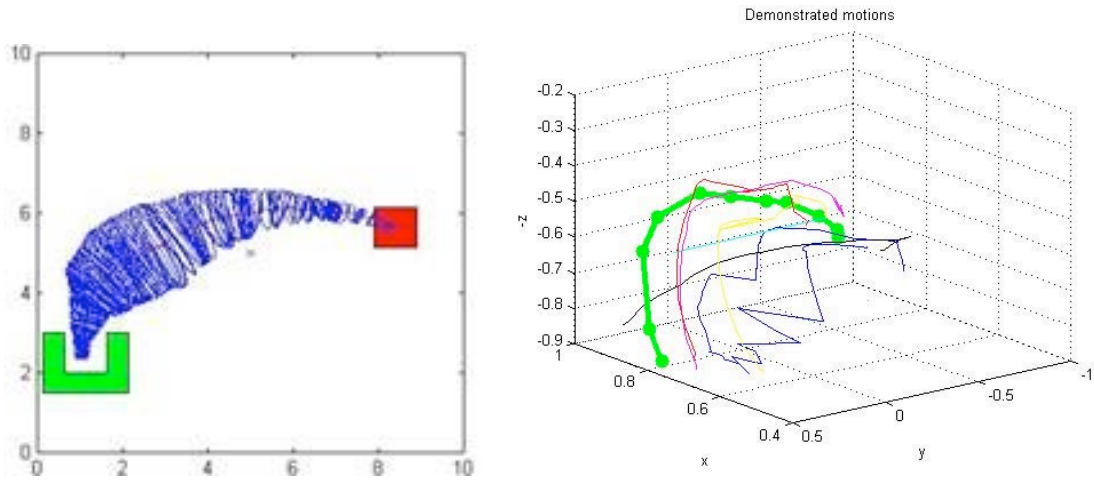
Figure 4. A sample autonomously generated flow tube for a "move object to bin" motion with previously unseen initial conditions (left), and the "move box to platform" ATHLETE limb 1 end-effector motion (right). Thin lines are demonstrated motions. Thick line is the trajectory of the learned motion applied to a new situation. Movement starts from the lower left and ends toward the upper right.

INTERPLANETARY COMMUNICATIONS

RECONFIGURABLE ROBOTIC SOFTWARE FOR RECOVERING FROM FAILURES IN LONG LIFE CYCLE MISSIONS

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1320

Vandana Verma, (PI), Mobility and Robotic Systems Section (348)
Ashley Stroupe, Mobility and Robotic Systems Section (348)
Gaurav S. Sukhatme, Department of Computer Science, University of Southern California
Nenad Medvidović, Department of Computer Science, University of Southern California
Jonathan Kelly, Department of Computer Science, University of Southern California
Hossein Tajalli, Department of Computer Science, University of Southern California

A.   OBJECTIVES

This project demonstrates advances in automatic sensor self-calibration, distributed failure detection, and automatic/continual assessment and reconfiguration of resources. For planetary exploration missions, human-in-the-loop rover sensor calibration and failure detection operations are time consuming and costly. We have developed a novel software architecture, adapted an accompanying lightweight component-based middleware platform, and created a series of software modules to optimally select both sensor and computing resources to meet task requirements, given individual sensor and processor failures, including automatic:

•   Self-(re)calibration of multiple visual and inertial sensors on a mobile rover.

•   Allocation of sensor/processor resources to achieve goals, given potential failures.

•   Model-driven run-time system adaptation with minimal performance overhead.

As an example, we demonstrate self-calibration of the camera-IMU transform for rover localization and mapping, dealing with IMU drift, unknown IMU orientation, lack of scale information in camera measurements, and lack of any prior knowledge of the environment. We present results from experiments with a monocular camera and a low-cost solid-state IMU, which demonstrate accurate estimation of the calibration parameters.

B.   APPROACH AND RESULTS

We developed an event-based, domain-specific architecture for robotic planetary rovers with multiple visual and inertial sensors. PLASMA uses an architecture description language (ADL) and a planning-as-model-checking technology to enable dynamic replanning. The ability to automatically generate adaptation plans based solely on ADL models and an application description simplifies the specification and use of adaptation mechanisms for system architects.

PLASMA Overview

PLASMA is based on the novel application of planning mechanisms to two different types of models: Application Domain Models and Adaptation Domain Models. Application Domain Models capture the possible states of application components and the actions that those components may perform. Each action modifies the application state in a defined way. Analogously, Adaptation Domain Models capture architectural states and actions: each state in the Adaptation Domain Model corresponds to a particular architectural configuration, and actions in the Adaptation Domain Model correspond to architectural modifications, such as the addition, removal, and replacement of components. Typical planning applies to application domains; PLASMA is unique in its usage of adaptation domains. Figures 1(a) and 1(b) are examples of the Application Domain Model. Figure 2 is an example of an Adaptation Domain Model.

PLASMA is a three-layer adaptive layered architecture (Figure 3). Components at a given layer monitor, manage, and adapt components at the layer below. Application-level components reside in the bottom Application Layer. The middle layer, called the Adaptation Layer, monitors, manages, and adapts components in the Application Layer. The Adaptation Layer is capable of modifying the Application Layer architecture to achieve a target application architecture, monitoring the Application Layer to ensure the target architecture is maintained, and making subsequent necessary modifications. The Planning Layer manages the Adaptation Layer and the generation of plans based on user goals and component specifications. Plans generated by the Planning Layer define both the target architecture for the Application Layer (the adaptation plan) and the actions to be carried out by the Application Layer (the application plan). Thus, the Planning Layer is capable of responding to changing system requirements or operational environments by regenerating plans. This three-layered architecture enforces a clear separation of concerns, whereby each layer in the system provides a different form of adaptation capability, and enables a high degree of autonomy.

In PLASMA, the only inputs provided by the system architect are (1) the application problem description, (2) component specifications written in an ADL, and (3) executable implementations of the components. The problem description consists of an initial state and a goal. The component ADL models are used by the ADL Model Parser components in the Planning Layer to generate Application and Adaptation Domain Model descriptions, as depicted in Figure 3. A domain model description presents a domain model in a standard language that is required for planning. Domain model descriptions and problem descriptions are provided to the Application Planner in the Planning Layer in Figure 3. Each of these two components generates a plan for one of the two bottommost layers: the Application Planner generates a plan for the Application Layer which specifies how to achieve the application goal, while the Adaptation Planner generates a plan for the Adaptation Layer that specifies how to arrive at the target application architecture. Application Plan and Adaptation Plan are executed in the Executor and Adaptation Analyzer in the two bottom-most layers.

PLASMA components are specified in ADL and passed to the ADL Model Parsers in the top layer in Figure 3. The ADL Model Parsers generate domain model descriptions that are passed to the appropriate Planner component. Component descriptions include (1) attributes,

(2) required and provided interfaces, and (3) pre- and post-conditions for each interface. Instead of explicitly (possibly rigidly) specifying the architecture, PLASMA infers the topology of the Application Layer from the plan and the component ADL models. Thus, when components are added, removed, or replaced, new plans can be automatically generated to achieve the goals. Changes to requirements represented in changes to ADL component model descriptions can thus be handled during runtime.

Planning

The Application Planner first finds an application plan that specifies how to achieve the goal from a given current state. The Application Planner then derives the target architecture of the Application Layer required to run that plan. The target architecture of the Application Layer, which is represented by a goal state in the Adaptation Domain Model, becomes the goal for the Adaptation Planner. The Adaptation Planner finds an adaptation plan that transforms a given current architecture of the Application Layer to the desired architecture. The application plan and adaptation plan are executed in Executor and Adaptation Analyzer components, respectively (see Figure 3).

Dynamic replanning can be initiated by the system architect (e.g., a new goal), or by unexpected changes to the Application Domain Model (a component becomes available or unavailable, e.g., due to a failure). Adaptation of the Application Layer occurs automatically whenever the adaptation goal or Adaptation Domain Model change, and is not preplanned.

Sensor Calibration

Visual and inertial sensors can be used to estimate motion if the six degrees-of-freedom transform between the sensors is accurately known. Calibration is a complex and time-consuming process, and must be repeated whenever the sensors are repositioned or significant mechanical stresses are applied. Inertial sensors are subject to drift and rely on accurately determining the gravity vector. Camera image measurements, unlike IMUs, reference the external environment and are therefore largely immune to drift. However, cameras are bearing-only sensors, which require both parallax and a known baseline to determine the absolute depths of landmarks. This baseline distance must be provided externally. The need for this complex calibration limits the versatility of integrated visual-inertial systems. We describe an algorithm, based on the unscented Kalman filter (UKF), for camera-IMU relative pose self-calibration using imperfect measurements from the sensors themselves to improve our estimates of related system parameters.

As part of the calibration procedure, we also simultaneously localize the camera-IMU platform and (if necessary) build a map of the environment. We formulate relative pose calibration as a filtering problem, and show that the sensor-to-sensor transform, the IMU gyroscope and accelerometer biases, the local gravity vector, and the metric scene structure can all be recovered from camera and IMU measurements alone. This is possible without any prior knowledge about the environment in which the robot is operating. We present experimental results from a monocular camera and low-cost solid-state IMU, which demonstrate accurate calibration. Target-based calibration involves sensor-related quantities only. The sensor state vector includes: position of the IMU in the world frame, unit quaternion orientation of the IMU frame with respect to the world frame, linear velocity of the IMU in the world frame, IMU

gyroscope and accelerometer biases, gravity vector in the world frame, position of the camera in the IMU frame, and unit quaternion orientation of the camera frame with respect to the IMU frame.

Adaptation: The Area Mapping Experiment

We simulated a scenario in which a rover searches and maps an area. To perform the tasks correctly, the robot's camera needs to remain calibrated. The simulated robot starts with a random battery level and the battery depletes as time passes, and the camera goes out of calibration at random times. The goal is to search and make a complete, accurate map of an area.

The simulation initially starts with the following components: Camera, IMU, FeatureExtractor, Filter, Mapper, Robot, and Search. Figure 1(a) shows the related domain model, which is extracted from the component models. Another component, Charger, adds an action to find and dock to a charging station. Charger can be added anytime, changing the Application Domain Model to the one shown in Figure 1(b), where the model calls for SearchForChargerAndDock action if BatteryIsLow=TRUE. Adding the Charger automatically results in replanning, creating a plan that makes use of the new component and actions.

The goal is to arrive at the MappingIsCompleted state. As the domain model in Figure 1(a) shows, the robot cannot perform the SearchForNewAreaToMap action when BatteryIsLow = TRUE. Consequently, the goal of the system cannot be achieved in such a state. PLASMA is able to calculate the application plan for this domain model in less than 1 second. The adaptation plan, which is calculated in around 3 seconds, sets the Application Layer's architecture.

Calibration Experiment

We used emulation to characterize calibration performance, with ground truth available. The simulated camera and IMU are modeled on the hardware in our laboratory. We used a set of 48 points on a simulated planar target as landmarks. To generate each simulated image, we projected the target points onto the image plane, and added uncorrelated, zero-mean Gaussian noise.

For the first simulation, initial translation and orientation errors were 5º along each axis; 2500 IMU measurements and 187 camera images (all 48 target points were visible) were processed over 25 seconds. The orientation estimates converge rapidly, with errors of < 0.1º in roll, pitch, and yaw. Translation errors decreased to 3 mm (± 7mm, 3-σ); see Figure 4.

For the second simulation trial, we used the same translation error as the TB-1 simulation, but with an initial orientation error of 10º along each axis. The orientation estimates again converged very rapidly, with residual errors of less than 0.1º in roll, pitch, and yaw after 25 seconds. The translation errors decreased to less than 3 mm ± 8 mm (3-σ) along each axis, within the 25-second simulation. These results indicate that the filter is able to accurately estimate the calibration parameters despite large initial errors in the relative camera orientation.

To verify the real-world performance of the calibration algorithm, we gathered actual image data of our (known) planar calibration target. The data included 15,025 IMU

measurements and 1,120 image frames (in which all target points were visible). Both the translation and orientation parameters converged to their steady-state values within 50 seconds. The final uncertainty for the parameter estimates was ~0.1º for orientation and <1 cm for translation. We evaluated calibration accuracy by comparing motion estimation results. We ran the algorithm twice: once with a hand-measured estimate of the camera-IMU transform and once with the calibrated estimate. The resulting errors, for both cases, are shown in Figure 4. Note that the RMS residual for the calibrated case is 3.08 pixels, which is less than half of the RMS value of 7.10 pixels for the hand-measured estimate. However, the gravity vector has a pitch of 0.9º relative to the z-axis of the calibration target. This misalignment was present despite efforts to orient the target exactly vertically (using a level), and indicates the importance of estimating the gravity vector accurately for calibration.

C.   SIGNIFICANCE OF RESULTS

Current robotic exploration missions are limited by their inability to autonomously adapt to sensor and/or processor addition and failure. Robustness to such changes will enable extended mission lifetimes, delivering more science over a longer operating period. Further, modular design and inherent adaptability will enable reuse across a variety of platforms. This work suggests the feasibility of self-calibrating, reconfigurable sensing systems for hostile, inaccessible environments.

PLASMA simplifies the specification and use of adaptation mechanisms for system architects by freeing them from designing the application architecture topology and plan for specific adaptations. The Adaptation Layer yields autonomy by using automatically generated plans to adapt the Application Layer. Placing the onus of automatic adaptation onto the Adaptation Layer reduces the burden of design on the architect because plans need not manually specify for each adaptation scenario. Moreover, PLASMA offers significant flexibility by allowing specification of adaptation requirements via ADL models of specific components in the Adaptation Layer.

The mapping experiment demonstrates how PLASMA-enabled applications transparently adapt to both foreseeable and unforeseeable conditions and requirements. Foreseeable conditions, such as going out of calibration, are built into application and adaptation plans. Thus, adaptations that handle these conditions are automatically realized without incurring the performance penalties of replanning. Unforeseeable conditions and requirements, such as an unexpected failure or the need to recharge, are addressed through architect-directed replanning. Unforeseen changes incur a penalty, but are typically rare and may improve overall mission performance (relative to failure or human-in-the-loop). Replanning times of <3 seconds are well within acceptable planning times for current planetary rovers. The ability to handle unforeseeable conditions and requirements sets PLASMA apart.

PLASMA is designed to simplify the task of specifying when and how the system should adapt. Automatically generated plans in our experiment ranged from 90 to 280 state-actions. Manually specifying policies of this size would be tedious and cumbersome for the architect. To add the recharging capability, for example, we only needed to specify one additional model along with the implementations of the one new component. Moreover, the system architect

maintains control over the process. The adaptations that result from new architect requirements are realized transparently, but the adaptations are only performed when the architect decides it is safe and appropriate to do so. Foreseeable adaptations/failures can be handled without architect intervention.

PLASMA has a number of benefits shown in the experiment. First, a policy-based approach can result in policies that conflict with each other or jointly make it impossible to achieve the goal. This cannot arise in PLASMA: if system constraints make the goal unachievable, PLASMA automatically notifies the architect before any adaptations are performed. Second, the adaptation plans preserve all software adaptation restrictions and requirements specified in the adaptation domain model. These requirements restrict architectural adaptation based on platform or application constraints. For example, when a component is removed, it will be disconnected from the rest of the architecture first. This avoids dangling ports, which can lead to communication failures.

In future, we plan to extend PLASMA to plan for temporary goals to address non-functional requirements, include contingent planning techniques [4] to reduce the frequency of replanning, and apply to distributed applications in the future, including addressing partial system failures.

D. FINANCIAL STATUS

Total funding for this task was $95,000, of which $90,380 has been expended.

E. REFERENCES

[1] N. Medvidovic and R. N. Taylor, "A Classification and Comparison Framework for Software Architecture Description Languages," *IEEE Trans. on Software Engineering* **26** (2000): pp. 70–93.

[2] R. Hermann and A. J. Krener, "Nonlinear Controllability and Observability," *IEEE Transactions on Automatic Control* **AC-22** (1977): pp. 728–740.

[3] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation," *Proceedings of the IEEE* **92** (2004): pp. 401–422.

[4] D. Shaparau, et al., "Contingent Planning with Goal Preferences, *21st National Conference on Artificial Intelligence*, pp. 927–934, AAAI Press, 2006.
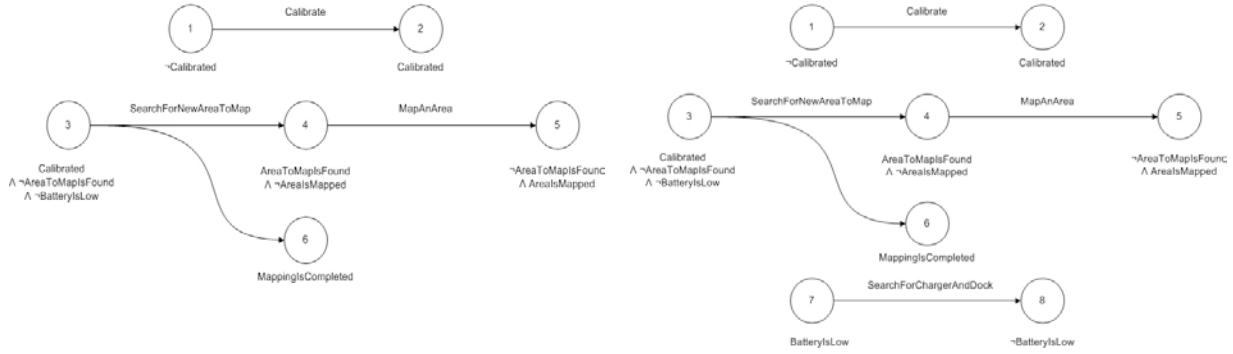
Figure 1. (a), left: Domain model for the mapping application; (b), right: Updated domain model.

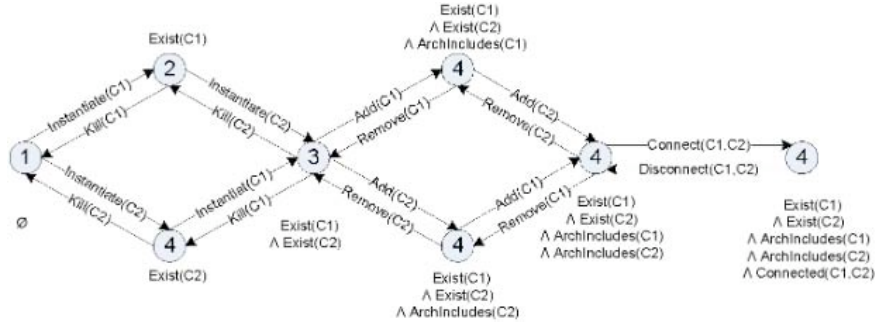**Figure 2.** An example Adaptation Domain Model.
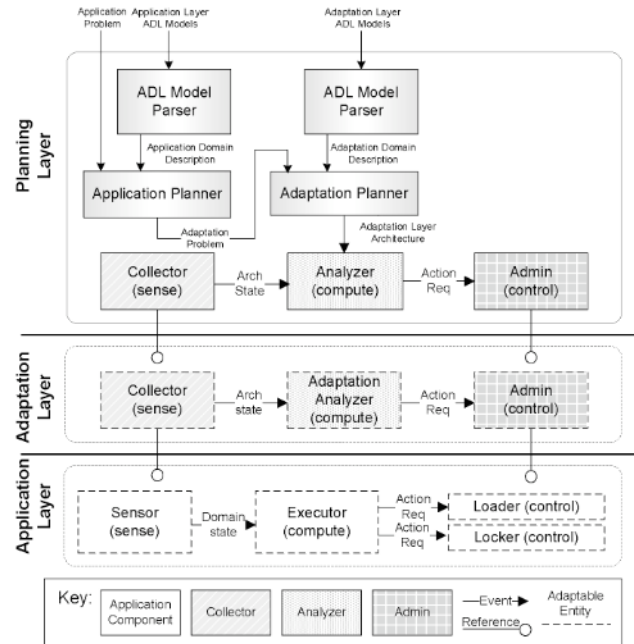
Figure 3. The PLASMA adaptive layered architecture.
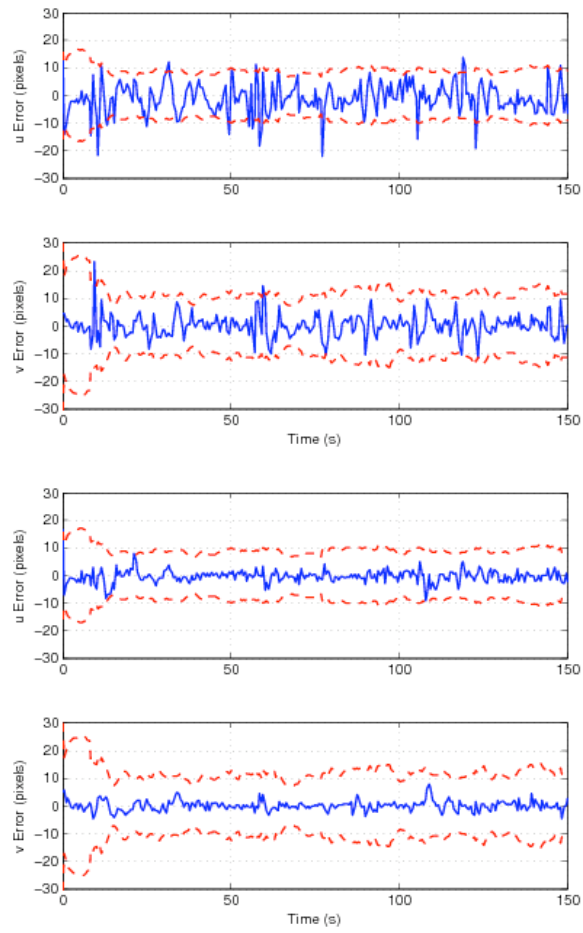
Figure 4. Residual pixel error and 3-sigma bounds for camera-IMU transform: manual (top) and calibrated (bottom).

**OPERATOR INTERFACE AND CONTROL SOFTWARE FOR THE
RECONFIGURABLE SURFACE SYSTEM TRI-ATHLETE**

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1324

Jeffrey S. Norris (PI), Planning and Execution Systems Section (317)
Daniela Rus, (Co-PI), Department of Electrical Engineering and
Computer Science, Massachusetts Institute of Technology
Marsette A. Vona, Department of Electrical Engineering and
Computer Science, Massachusetts Institute of Technology

A.   OBJECTIVES

The objective of this project was to develop graphical operator interface methods for modular reconfigurable articulated surface systems in general, and a specific instantiation thereof for JPL's Tri-ATHLETE [1].

Reconfigurable modular hardware such as Tri-ATHLETE (Figure 1) promises to greatly extend the capability of future exploration missions for a relatively small additional cost. Whereas existing missions based on monolithic hardware can only perform a limited set of predefined operations, modular hardware can potentially be reconnected and recombined to serve a range of functions. In the case of Tri-ATHLETE, these include: lifting and carrying specific payloads; environment manipulation (digging, pushing, drilling); several locomotion modalities; and various crew habitat configurations. Modular hardware also promises to enable graded reduction in mission capabilities in the event of hardware faults — or even total repair when surplus modules are available — and the multipurpose nature of individual modules can potentially be leveraged for technology re-use across missions.

The full realization of these promises is contingent not just on the development of the hardware itself, but also upon the availability of corresponding software systems with algorithms that enable operators to rapidly specify, visualize, simulate, and control (1) particular assemblies of modules, and in the case of articulated reconnectable hardware like Tri-ATHLETE, also (2) feasible motions of the assembly, and (3) disconnect/reconnect actions that change assembly topology. Existing operations software, such as JPL's Ensemble [2], has been architected for monolithic (nonmodular) robots and does not include these capabilities. Conversely, user interface software previously reported for modular robots has typically been tightly connected with the particular system for which it was developed [3–6].

B.   APPROACH AND RESULTS

The core advancement of this work was the development of a novel set of algorithms for dynamically maintaining a reduced coordinate model of any connected assembly of robot modules. The kinematics of individual modules are first modeled using a catalog of 12 standard 3D robot joints (this modeling step need be done only once). Then, individual modules can be assembled into any closed- or open-chain topology. The system automatically maintains a spanning tree of the overall configuration, which ensures both efficiency and accuracy of the onscreen representation.

We implemented these algorithms in the context of a full 3D interactive kinematic simulator [7]. Both the algorithms and the implementation are generic across a broad class of modular robots; we tested them with use-cases both for a specific model of Tri-ATHLETE and also for "CKBot," a completely different modular robot taken from the current literature [6].

A canonical use-case for Tri-ATHLETE is assembling two mobility modules onto a pallet. Figure 3 shows the evolution of the robot model from an initial state with two distinct modules — one Tri-ATHLETE and one pallet — into a connected configuration. The process is equivalent to adding the second Tri-ATHLETE module, resulting in the full three-module assembly shown at the top of the figure.

Zooming in to consider the intermodule attachment process in more detail, there are three key transitions. First, the user specifies that there is a rigid connection between the two modules. Second, the system automatically applies a least-squares numeric solver to move the modules together (Figure 3C). Third, the connection constraint is effectively converted into an implicit form (Figure 3E) for all subsequent operations.

The operator may interact with the model at any time to specify kinematic motions; Figure 2 shows three motions that an operator might specify for the Tri-ATHLETE-pallet-Tri-ATHLETE assembly. Motion specification can be as straightforward and intuitive as clicking on any part of the model and dragging with the mouse; the system will enforce all necessary joint constraints. Additional constraints can also be defined using virtual joints and links [8].

Our system is, of course, not limited to just the above canonical assembly. We tested the generality of our algorithms in two ways: first, we constructed and virtually operated a novel Tri-ATHLETE configuration (Figure 5); second, we modeled the CKBot modules, constructed a novel assembly of them, and virtually operated it (Figure 4).

C.   SIGNIFICANCE OF RESULTS

Current operator interfaces, both at JPL and in the broader exploration robotics community, are largely focused on nonreconfigurable hardware. Modular and reconfigurable articulated surface systems such as Tri-ATHLETE are a new and promising direction to increase mission function and reliability while managing cost. But until now, JPL has used generic CAD, simulation, and animation tools as a stand-in for a true modular robot operator interface. This workflow is extremely time-consuming, and is not suited for use in an operations context.

The algorithms we have developed can form the basis for a new kind of interface, visualization, and simulation tool that is specifically suited to modular robots. The resulting workflow can be very rapid; for example, a novel Tri-ATHLETE module assembly was designed and virtually operated with a prototype of our system in only a few minutes (Figure 5).

Also, unlike other hardware-specific modular robot user interfaces systems that have been reported in the literature, our algorithms are not specific to Tri-ATHLETE. They are based on a generic approach to modeling closed- and open-chain modular robot kinematics, and hence will be re-usable across a broad class of modular systems.

The algorithms we have developed can thus fill the previously open gap between the low-level tedium of general-purpose tools and the inflexibility of existing hardware-specific systems.

D.   FINANCIAL STATUS

The total funding for this task was $100,000, all of which has been expended.

E.   ACKNOWLEDGEMENTS

The Tri-ATHLETE model shown in figures was developed by Mark Powell at JPL.

F.   PUBLICATIONS AND PRESENTATIONS

[A]   Marsette Vona, "A Graphical Operations Interfaces for Modular Surface Systems," *Proceedings of the IEEE Aerospace Conference*, Big Sky, Montana, March 2010.

G.   REFERENCES

[1]   B. H. Wilcox, "ATHLETE: A Cargo and Habitat Transporter for the Moon," *Proceedings of the IEEE Aerospace Conference*, Big Sky, Montana, 2009.

[2]   Jeffrey Norris, Mark Powell, Jason Fox, Kenneth Rabe, Hsiang Shu, Michael McCurdy, and Alonso Vera, "Ensemble: an Architecture for Mission-Operations Software," NASA NPO-41814, 2008.

[3]   T. Troncy, M. T. Martinez, S. E. Baba, and C. Hugues, "Modular Robots—Graphical Interactive Programming," *Proceedings of the IEEE International Conference on Robotics and Automation*, Philadelphia, Pennsylvania, April 24–29, 1988: pp. 1739–1742.

[4]   P. C. Y. Chen, I.-M. Chen, I.-G. Kang, W. Chen, and G. Yang, "Development of a Simulation and Control Environment for Modular Robotic Systems," *Proceedings of the International Conference on Control, Automation, Robotics and Vision*, Singapore, December 8–11, 1998, pp. 504–508.

[5]   M. Bordignon, K. Støy, D. J. Christensen, and U. P. Schultz, "Towards interactive programming of modular robots," *Proceedings of the IROS Workshop on Self-Reconfigurable Robots, Systems and Applications*, Nice, France, September 22, 2008.

[6]   M. Yim, P. White, M. Park, and J. Sastra, "Modular Self-Reconfigurable Robots," *Encyclopedia of Complexity and Systems Science*, Springer, 2009, pp. 5618–5631.

[7]   Marsette A.Vona, III, "Virtual Articulation and Kinematic Abstraction in Robotics," Ph.D. dissertation, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, August 2009.

[8]   M. Vona, D. S. Mittman, J. S. Norris, and D. Rus, "Using Virtual Articulations to Operate High-DoF Manipulation and Inspection Motions," *Proceedings of the Field and Service Robotics Conference*, Cambridge, Massachusetts, July 2009.

H.   FIGURES
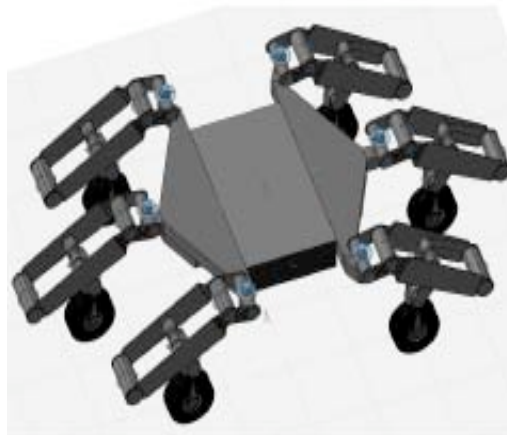
Figure 1. Rendering (from our interface system software) of two Tri-ATHLETE modules attached to a central rectangular payload pallet.
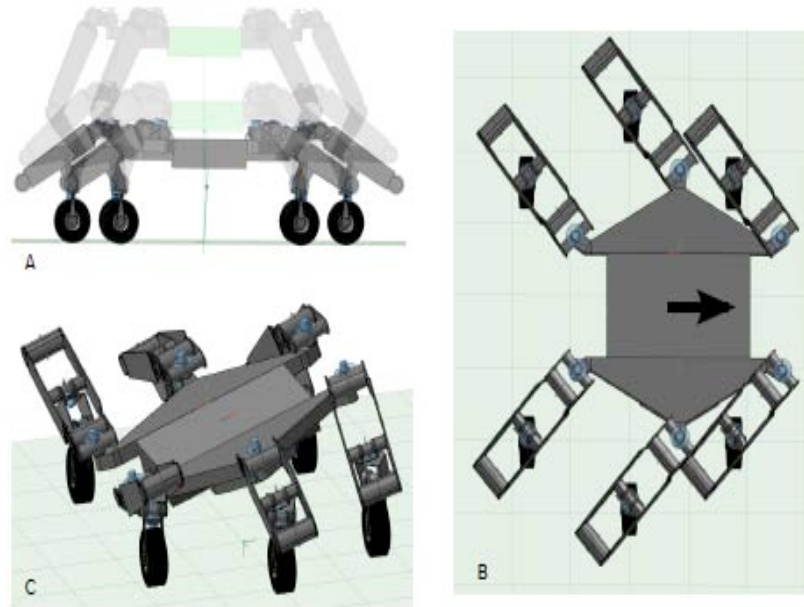
Figure 2. Any assembly of kinematic modules may be directly operated in our system by click-and-drag direct manipulation. Here the canonical configuration of two Tri-ATHLETE modules and one pallet is operated in lifting (A), sliding (B), and tilting (C) motions.
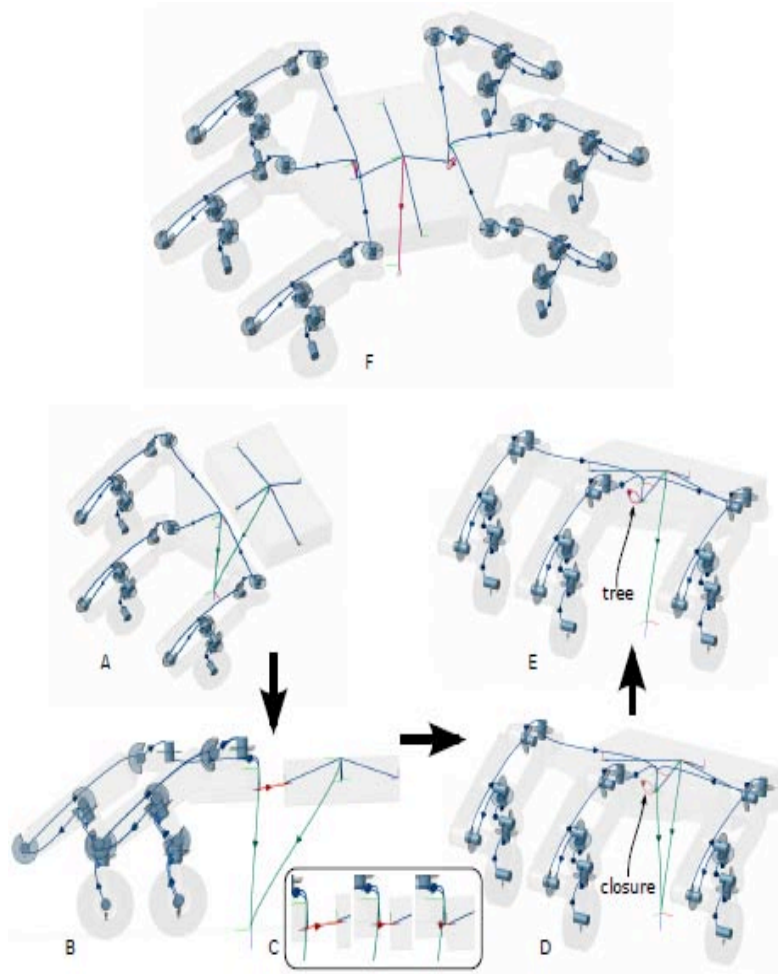
Figure 3. Model evolution as one Tri-ATHLETE module is attached to a pallet module (Parts A to E). Part F shows the final result after attaching two Tri-ATHLETEs to the pallet.

Figure 4. An example with the "CKBot" system from the modular robotics literature [6]: (A) a graphical model of a 12-module quadruped walker is shown in our interactive interface. It took only about 2 hours to design CKBot module models from scratch. Virtually assembling the walker then took about 10 minutes. The modules were assembled graphically without any hand-coding. As before, the resulting model can be operated with click-and-drag direct manipulation, e.g., to design poses for a walking sequence (C, D).
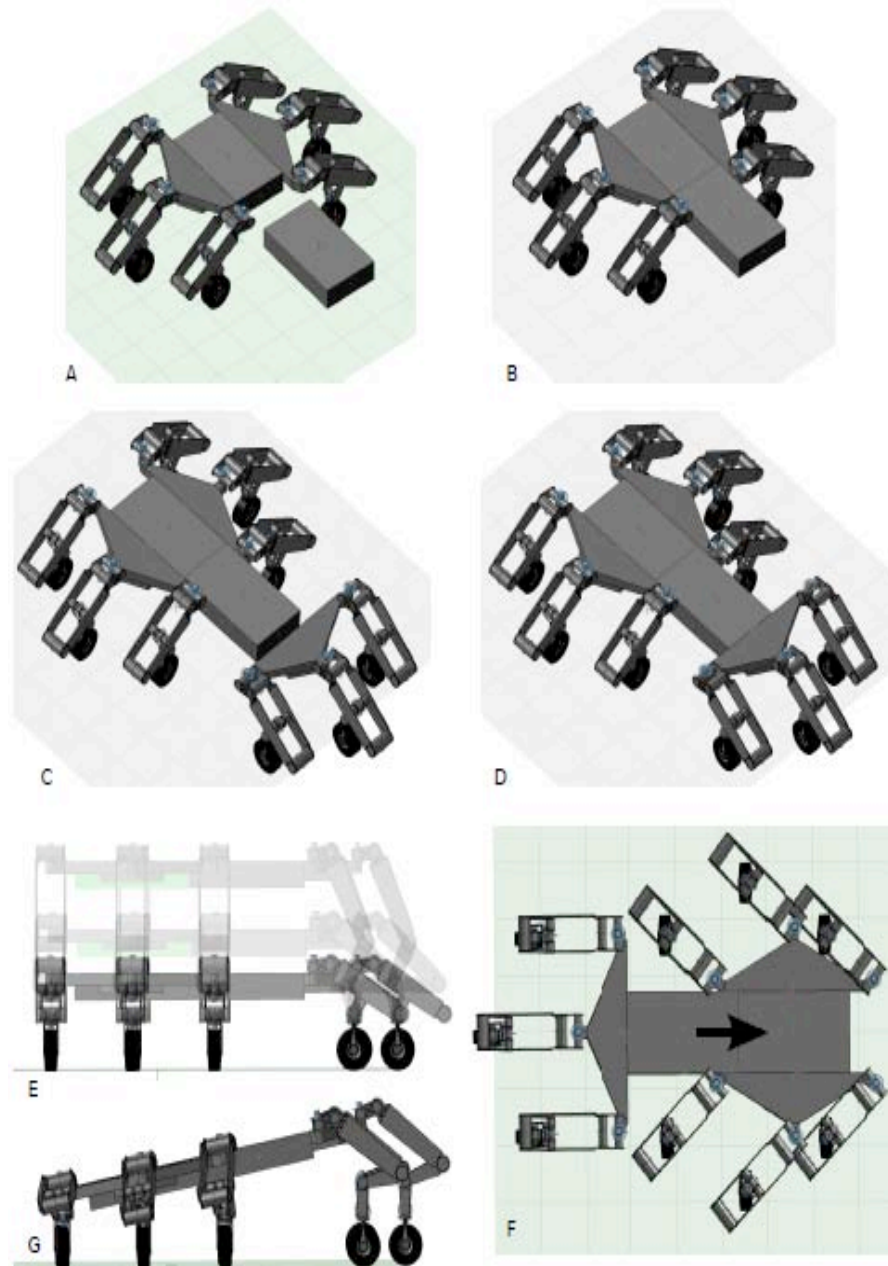
Figure 5. Once the individual Tri-ATHLETE modules have been modeled, new mission scenarios can be rapidly explored. In this example, a novel structure is created by adding a second pallet and a third Tri-ATHLETE module (A-D) to the standard configuration (Figure 1). It took only a few minutes to create this novel assembly of modules, and the generic click-and-drag direct manipulation interface allows operation of similar lift (E), slide (F), and tilt (G) motions as for the standard single-pallet assembly (Figure 2).

NEXT GENERATION LEADERS AND INNOVATORS

ATHLETE LUNAR ROVER CONTROL FOR CHALLENGING TERRAIN

Jaret B. Matthews (PI), Mobility and Robotics Systems Section (347)
Brian Wilcox (Co-PI), Autonomous Systems Division (340)
David S. Wettergreen (Co-PI), Robotics Institute, Carnegie Mellon University
Scott J. Moreland, Mechanical Engineering Department, Carnegie Mellon University

A.   OBJECTIVES

The objective of this work was to achieve significant increases in mobility for the ATHLETE rover platform. Specifically, the goal was to attain the ability to continuously ascend slopes of 20 degrees of loose material by means of detecting slip on board the rover and using a controller to adjust vehicle angle of attack (heading with respect to incline) and/or adjust individual wheel speeds for traction control.

This work built on a previous SURP award that investigated nominal terrain efficiency and traction characterization of the ATHLETE rover's many locomotion modes. The current work extended to more extreme terrain and attempts to make advances in sensing and control on steep slopes.

The study had two distinct tasks. Initially, an investigation of vehicle response to terrain in a relevant setting would be used to design the controller and method of sensing slip. The second stage is to develop and demonstrate reactive control based on the instantaneous slip measurement to keep the vehicle in low slip regime while ascending steep slopes. The development of a small-scale ATHLETE (SSA) rover was undertaken to support the mobility experimentation and control development phases. This rover would be used at the Glenn Research Center (GRC) Simulated Lunar Operations (SLOPE) Laboratory to assess the vehicle response and slip/traction controller in a well controlled simulated lunar terrain.

B.   APPROACH AND RESULTS

1.   Development of the SSA

The development of a controller for slope climbing required the ability to accurately reproduce terrain and measure vehicle response. This could only be undertaken in a laboratory environment, so the GRC SLOPE facility was chosen for the experimentation. A small-scale version of the ATHLETE vehicle (Figure 1) was developed that could be used at the GRC facility. This system has 18 degrees of freedom and will eventually be upgraded to 24 degrees of freedom. The platform is able to perform multiple mobility modes such as 4- or 6-

wheel rolling, thrust-slide (inching), and 6 d.o.f center of gravity control, and will soon be capable of walking. In addition to a kinematic similarity to the larger ATHLETE prototypes, the small-scale version shares the same avionics hardware and software. This allows for the SSA to be used as a software development platform.

The SSA has a 48" × 48" footprint, 10" wheels, and configurable 150–250 lb vehicle weight. This size makes the system ideal for use in the GRC SLOPE facility (Figure 2). Initial study of the vehicle response and performance to slope climbing is underway. This current investigation is primarily assessing the ability to control vehicle angle of attack (heading with respect to incline) to reduce slip and power number (energy consumption for a task) for steep slope climbing. Following this work, a controller based on instantaneous slip measurement will be developed during this SURP grant to demonstrate this example of slope climb autonomy.

2.   Control Development

The control techniques are currently being developed and have yet to be implemented on the SSA. A method of maintaining slip within a threshold while maximizing power number will be applied. Control of the vehicle drive arc while using slip measurement as the controller input will aid in climbing slopes under low slip and high efficiency. Preliminary work at the GRC SLOPE facility has shown significant gains in vehicle safety (slip/traction) and energetics (power number) (Figure 3 and Figure 4).

C.   SIGNIFICANCE OF RESULTS

The application of recent investigations into slope climbing for wheeled vehicles on the Moon will provide an advance for planetary mobility. Many of these lessons learned and methods to evaluate performance will be applied to the proposed work with ATHLETE. These techniques reduce the wheel thrust required, leading to higher traction, and allow for slip to be controlled and maintained within a threshold. Combining this with the high degree-of-freedom ATHLETE system will allow for significant advances in performance in difficult and extreme terrain.

The development of vehicle test apparatus and procedure now allows for ongoing evaluation of the ATHLETE lunar rover and follow-on research. This can be undertaken in the GRC SLOPE laboratory or at annual analog site field trials. Furthermore, the ability to utilize the mobility system as a platform for general vehicle experimentation is now possible. The high d.o.f chassis can be reconfigured to mimic many mobility platform types, and comparative studies can now be conducted with ease in order to analyze various locomotion modes (walking, rolling, hybrid braked/roll).

In addition to technical gains, the proposed work has resulted in strengthened ties between Carnegie Mellon University's Robotics Institute and JPL's Robotic Hardware Systems group. The SURP award has also provided an avenue for valuable educational and practical experience for students directly involved with this effort.

D.   NEW TECHNOLOGY

No new technology was developed.

E.   FINANCIAL STATUS

F.   ACKNOWLEDGEMENTS

G.   PUBLICATIONS

None.

H.   REFERENCES

None.

I.   FIGURES

Figure 1. The small-scale ATHLETE (SSA) rover is suitable for a wide range of mobility experiments. This system was developed by the involved student during the 2010 SURP grant to support mobility experiments and software development for traction control on steep slopes.
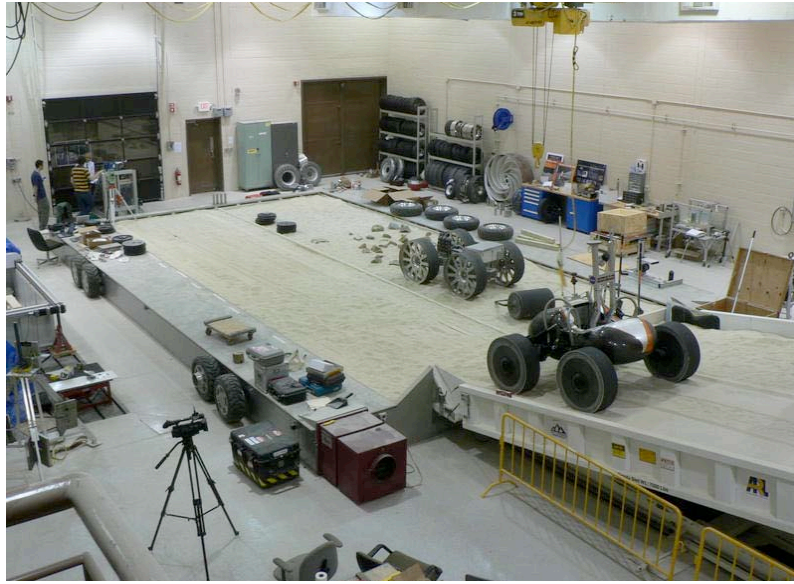
Figure 2. The GRC Simulated Lunar Operations (SLOPE) laboratory where the SSA mobility experimentation will occur. It features a 60' × 20' sandbox with multiple simulant types, hydraulic adjustable tilt bed from 0–48 degrees, and Drawbar pull rig for traction characterization.
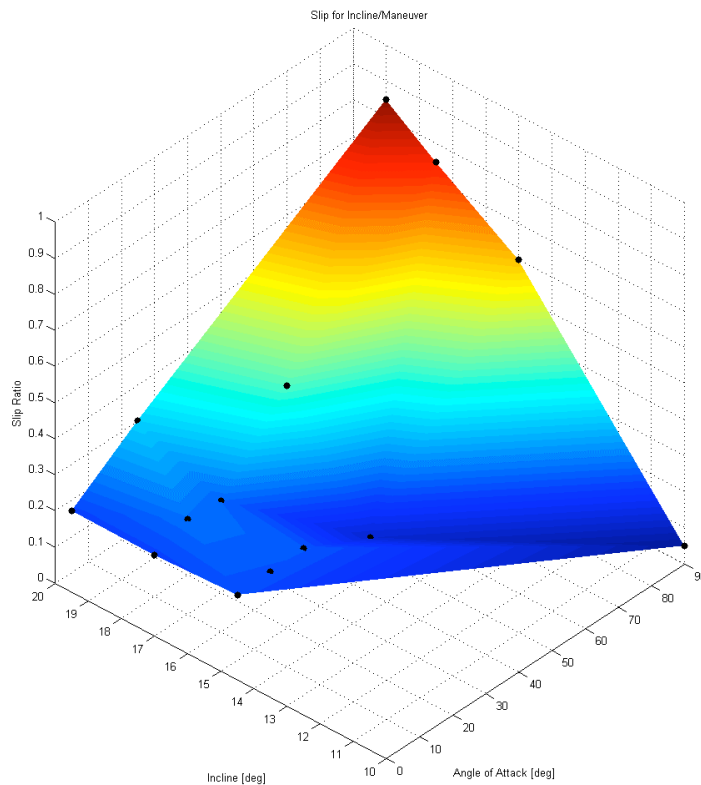
Figure 3. Slope climb experimental results displaying slip metric for range of incline angles and angles of attack (black "dots" are actual data points).
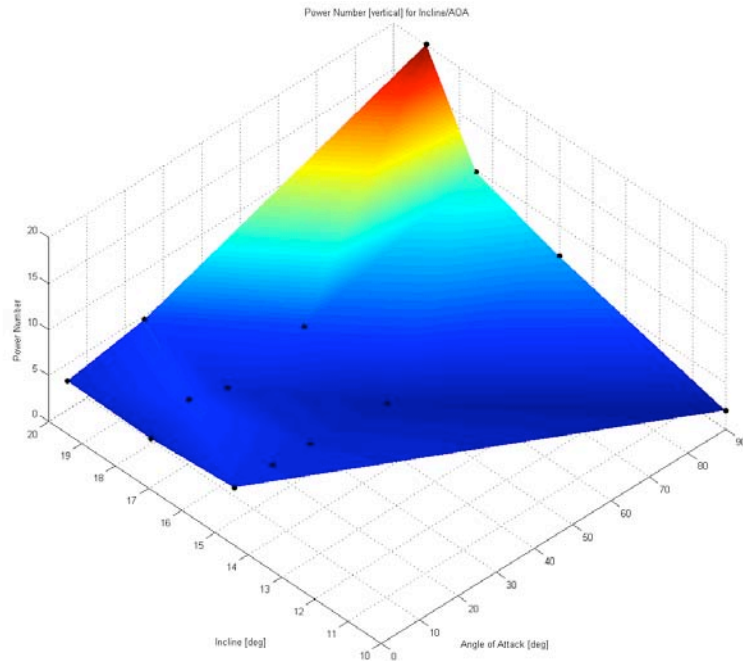
Figure 4. Slope climb experimental results displaying power number metric for range of incline angles and angles of attack (black "dots" are actual data points).

EXOPLANET MISSION DESIGN: LEVERAGING PRINCETON STUDENTS AND TEAM X EXPERTISE FOR INNOVATION AND EDUCATION

Jeffrey Thomas Booth (PI), Mission Systems Concepts Section (312)
Keith Warfield, Mission Systems Concepts Section (312)
Jeremy Kasdin, Department of Mechanical Engineering, Princeton University

A.   OBJECTIVES

The objective of the SURP proposal was to engage the next generation of leaders and innovators in aerospace engineering and exoplanet discovery and introduce them to JPL and the concurrent engineering environment of Team X. Led by Professor Jeremy Kasdin of Princeton, an expert in exoplanet detection techniques and PI of the JPL-partnered Telescope for Habitable Earth and Interstellar/Intergalactic Astronomy (THEIA) mission concept, the Princeton Space Systems Design students focused on an innovative mission design concept during their spring undergraduate course. Their concept, Doomsat, was a technology testbed of deformable mirrors on a small space platform. At the end of the course, students were flown to JPL to participate in a Team X study/review of Doomsat.

B.   APPROACH AND RESULTS

Students, along with the Teaching Assistant for the course and Prof. Kasdin, developed the Doomsat mission concept throughout the academic course. JPL provided the standard Team X briefing package requirements to the students in advance for them to fill out and present. Students spent three days at JPL. On Day 1, the students presented for the briefing package to Team X management (modeling our typical "pre-session" briefing") and leading engineers from the Mission Systems Concepts Section (312). Particular issues or concerns were raised, and the students had one day to turn around and adapt their design before the Team X session began in earnest. Days 2 and 3 were fully engaged in a Team X Red Team Review, where the students again presented their mission in abbreviated form, then worked in the concurrent engineering environment with the JPL Team X engineers to understand the trade space around their point design for Doomsat.

The technology tested in the Doomsat design is deformable mirrors (DM), which are a critical component of nearly any future instrumentation for an exoplanet mission. In order to control the wavefront at the stability required to separate a single planetary photon from the billion stellar photons, the DMs must perform flawlessly. They are not yet flight qualified, and the student project would flight qualify DMs in a low-Earth orbit smallsat with a small aperture telescope and camera.

Students were paired by their subsystem responsibilities in the class to their corresponding Team X engineering subsystem leads (two to three students per chair). Over two days, the JPL engineers worked with the students to confirm their overall design and point out areas for improving margins or considering implications for design trades, fully engaging the students in the learning/concurrent engineering process.

In follow-up activity from the Team X session on Doomsat, students at Princeton have continued to develop laboratory hardware testbeds in preparation for future flight opportunities.

After each day's Team X session, Lab tours and meetings with JPL engineers were arranged, including a briefing with the JPL Chief Engineer and interviews with the JPL PlanetQuest website.

C.   SIGNIFICANCE OF RESULTS

Princeton does not have their own concurrent engineering design capability; hence, at the end of the Team X collaboration, students had a clearer understanding of the mission design relationships among science instruments, cost, and schedule, and the trade-offs necessary to stay within cost and schedule while preserving the opportunity to acquire high-quality science.

This engagement has opened a recruiting pipeline for JPL as well, and several Princeton students from the SURP activity are currently undergoing the recruiting/interviewing process at JPL.

Overall, the students were excited by the interactions at JPL, informed by the way JPL approaches systems engineering of space missions, and the flexibility and creativity afforded by the concurrent engineering process in Team X. Feedback also provided useful input on hardware testbeds being developed back in Princeton.

D.   FINANCIAL STATUS

The total funding for this task was $25,000, all of which has been expended

E.   ACKNOWLEDGEMENTS

We would like to acknowledge Paula Grunthaner and the SURP program. We would also like to thank Rebecca Wheeler and the rest of the Team X management and engineering chairs for their support and engagement with the students.

F.   PUBLICATIONS

None.

G.   REFERENCES
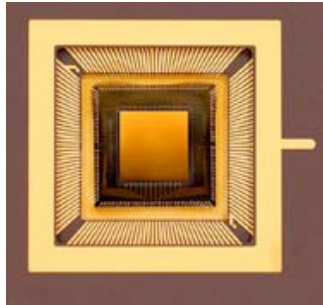
None.

H.   FIGURES

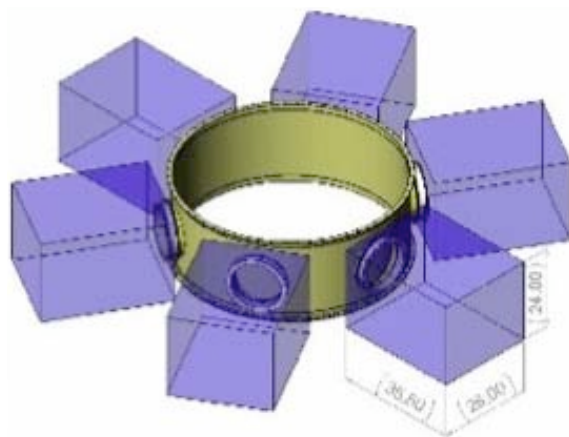**Figure 1.** A deformable mirror, which was designed to be tested on flight by Doomsat.

Figure 2. The ESPA ring is a standard interface for ride-share satellites. Doomsat would couple to one of the openings in the ESPA ring.

RECONFIGURABLE, REUSABLE MISSIONS

ADAPTIVE ROVER NAVIGATION

Michael McHenry (PI), Mobility and Robotic Systems Section (347)
Thomas Howard, Mobility and Robotic Systems Section (347)
Mihail Pivtoraiko, Robotics Institute, Carnegie Mellon University
Alonzo Kelly, Robotics Institute, Carnegie Mellon University

A.   OBJECTIVES

This task focused on the research and development of autonomous navigation of planetary rovers in rough terrain. The task satisfied its process and technical objectives. On the process side, the participating students received a year-long practical experience in key elements of software systems design and delivery, and in working in a team environment. On the technical side, the results of this work provided a measurable improvement in rover navigation over the state-of-the-art approach currently in use in the MER Spirit and Opportunity rovers. The average distance traveled to the goal was decreased, while the probability of successfully reaching the goal was increased by modifying the navigation system's selection of motion alternatives. The minor added computational cost is more than offset by the ability to consider fewer prospective motions. This report describes the quantitative improvement of rover navigation that was attained and details the developed general principles of generating good motion alternatives for navigation.

B.   APPROACH AND RESULTS

1.   Measures of Performance

In order to orient the development of improved rover navigation, the performance of the state-of-the-art approach to rover navigation was studied and evaluated. The GESTALT navigation system [1,2], currently utilized by the Mars Exploration Rovers (MER), was chosen as a representative solution for modern autonomous planetary navigation. In order to evaluate new approaches to rover navigation, a system of rigorous and quantitative measurement of navigation quality was established.

To devise this system, we utilized our previous experience in field experimentation with rover navigation using research rover prototypes in rough terrain at the JPL Mars Yard. From that experience, the quantitative measures of navigation quality were developed. Two relevant measures that were best suited for comparing navigation approaches were outlined as success or failure to reach the goal and computation runtime.

These quantities are well suited for evaluating navigation because they are simple to measure directly, well defined in the context of autonomous rover navigation, and reflect the practical concerns of rover autonomy for planetary exploration. The software was set up to record the values of these measures during navigation in a consistent manner, in order to make comparison meaningful. The first measure (success in reaching the goal) was a Boolean value (whose distribution can be represented as probability over a number of experiments), while SI units were used for other measures. Energy expenditure of the actual rover was derived from the measurement of voltages that were applied to the rover's motors during driving, while in simulation an estimate of it was produced using available data. The estimate was calibrated by comparing with the data recorded using the actual rover.

2.   Motion Adaptation for Rover Navigation

Once the relevant measures of navigation performance were developed, they were used to compare alternative approaches to rover navigation. All alternatives shared the same general architecture as GESTALT. This architecture involves a planning component, responsible for computing rover motion, which functions in tandem with other components, including state estimation and perception. In GESTALT, a local planner evaluates a set of motion alternatives and suggests the best one to follow, and a global planner suggests the general direction toward the goal. The achievement of this work is the development of an improved navigator design solely by modifying the local motion selection component of the system. While GESTALT considers a fixed set of constant curvature arcs, this constraint was relaxed in this work. In particular, we found notable performance improvement by endowing motion alternatives with a capacity to adapt to the sensed environment around the rover.

Finding the sets of motion primitives (paths) that improve the relevant measures of navigation has been the focus of this work. A number of general approaches to primitive design have been developed. It was verified experimentally that solely by replacing the constant curvature arcs in GESTALT's local planner with a set of arcs whose curvatures were modified via an optimization framework [3] as discussed below, it is possible to achieve a significant improvement of navigation performance. The developed approach can be configured to incur no additional cost in runtime or memory storage since a fewer number of adapted arcs can perform on par with a significantly larger number of fixed arcs.

In addition to the adaptation of constant curvature arcs, more general forms of motion adaption were investigated. In particular, optimization of non-constant curvature motions (clothoids) was implemented and various associated constraints were considered. Specifically, we examined the changes to performance resulting from 1) mixing constant curvature-arcs with non-constant curvature arcs; 2) allowing the end-points of the non-constant curvature arcs to differ from the original constant-curvature baseline set; and 3) allowing the end vehicle orientation to differ from the original baseline set. Based on those experiments, our initial conclusion is that the incremental benefits beyond that of constant-curvature arc adaptation was not justified given the additional computational overhead. If in the future, planetary rover navigation is less computation constrained, it may be worth revisiting the potential benefits of non-constant curvature arcs, particularly since previous work by [4,5] has shown the importance of utilizing the vehicle's maximum curvature for efficient maneuvering.

The arc adaptation algorithm and the result of its operation are presented in Figures 1 and 2, respectively. The results using the adapted arc sets have been compared to the baseline arc set as utilized in the standard GESTALT system. The experiments were focused on estimating both measures of quality outlined above: computation time required in planning and probability of finding a path amidst dense obstacles. The former measure is presented in Figure 3a. Two types of navigation are compared: standard set of GESTALT arcs (consisting of 23 fixed arcs; blue line) and a set of only 7 arcs, endowed with adaptation capacity (red line). Despite the additional computation overhead of adaptation, by virtue of considering only 7 arcs, the adaptation-enabled navigation consistently requires less computation. Furthermore, Figure 3b demonstrates the results in attempting to navigate in dense obstacle environments and focuses on estimating the probably of failure to find a path to follow. The rover was placed inside one of the three cul-de-sacs (subfigure c). The navigation goal was placed outside of cul-de-sacs, and the rover was required to maneuver out. Similar to Figure 3a, the blue line represents the standard GESTALT arc set, while the red line represents a set of 7 arcs endowed with adaptation capacity. The latter was able to guide the robot out of all cul-de-sacs, while the former was not able to compute the required motion in a significant percentage (10%–60%) of trials.

C.   SIGNIFICANCE OF RESULTS

This task focused on improving autonomous rover navigation by redesigning an important component of conventional rover navigation systems, the process of selecting the immediate motions of the rover. Through no other change but the choice of locally relaxing motion alternatives considered by this process, a significant performance improvement was attained. Relevant performance measures have been developed that allowed to quantify the benefit of this contribution. Throughout this analysis, it was determined that the enhanced rover navigation system, developed by this task, offers visible advantages over the state-of-the-art navigation solutions.

The software produced in this work has been integrated with the CLARAty [6] software architecture. Thus, this task has successfully met its process and technical objectives. In the future, we plan to seek opportunities for infusion of the technology into the upcoming rover missions in order to enable future rover systems to benefit from this research.

Furthermore, we are happy to report that our former student collaborator, Thomas Howard, joined JPL after his graduation last August.

D.   NEW TECHNOLOGY

The new technology that resulted from this research task is a novel component of a state-of-the-art autonomous rover navigation solution. This component allows a significant measurable improvement in efficiency and robustness of autonomous rover navigation.

E. FINANCIAL STATUS

F. ACKNOWLEDGEMENTS

G. PUBLICATIONS AND PRESENTATIONS

[A]   Mihail Pivtoraiko and Alonzo Kelly, "Fast and Feasible Deliberative Motion Planner for Dynamic Environments," in *Proc. of the International Conference on Robotics and Automation*, Kobe, Japan, May 12–17, 2009.

[B]   Mihail Pivtoraiko, Alonzo Kelly, and Issa Nesnas, "Autonomous Robot Navigation using Advanced Motion Primitives," in *Proc. of the IEEE Aerospace Conference*, Big Sky, Montana, March 7–14, 2009.

H. REFERENCES

[1]   Steven Goldberg, Mark Maimone, and Larry Matthies, "Stereo Vision and Rover Navigation Software for Planetary Exploration," in *Proc. of the IEEE Aerospace Conference*, Big Sky, Montana, March 9–16, 2002.

[2]   Mark Maimone, Chris Leger, and Jeffrey Biesiadecki, "Overview of the Mars Exploration Rovers' Autonomous Mobility and Vision Capabilities," in *Proc. of the IEEE Int. Conf. on Robotics and Automation*, Roma, Italy, April 10–14, 2007.

[3]   Thomas Howard, "Adaptive Model-Predictive Motion Planning for Navigation in Complex Environments," Ph.D. Thesis, Carnegie Mellon University, 2009.

[4]   J. Barraquand and J.-C. Latombe, "On Nonholonomic Mobile Robots and Optimal Maneuvering," *Proc. of the IEEE International Symposium on Intelligent Control*, Albany, New York, September 25–26, 1989.

[5]   J. A. Reeds and L. A. Shepp, "Optimal Paths for a Car that Goes Both Forwards and Backwards," *Pacific Journal of Mathematics* **145** (1990): pp. 367–393.

[6]   Issa Nesnas, "CLARAty: A Collaborative Software for Advancing Robotic Technologies," NASA Science Technology Conference, University of Maryland, June 19–21, 2007.

I.   FIGURES

Input: curvature of arc $\kappa \in [\kappa_{min}, \kappa_{max}]$
Output: curvature of relaxed arc $\kappa_{rel} \in [\kappa_{min}, \kappa_{max}]$

if $\text{cost}(\kappa - \Delta\kappa) < \text{cost}(\kappa)$ AND $\text{cost}(\kappa - \Delta\kappa) < \text{cost}(\kappa + \Delta\kappa)$ then
$\quad\quad \Delta\kappa \leftarrow -\Delta\kappa$
else if $\text{cost}(\kappa - \Delta\kappa) >= \text{cost}(\kappa)$ AND $\text{cost}(\kappa + \Delta\kappa) >= \text{cost}(\kappa)$
then
$\quad\quad \kappa_{rel} \leftarrow \kappa$;
$\quad\quad$ return κ_{rel};
$continue \leftarrow$ true;
$cost\_1 \leftarrow \text{cost}(\kappa)$;
$cost\_2 \leftarrow cost\_1$;
$\kappa_{temp} \leftarrow \kappa$;
while $|\Delta\kappa| > \Delta\kappa_{min}$ do
$\quad\quad$ while $continue$ AND $\kappa + \Delta\kappa >= \kappa_{min}$ AND $\kappa + \Delta\kappa <= \kappa_{max}$ do
$\quad\quad\quad\quad continue \leftarrow$ false;
$\quad\quad\quad\quad cost\_1 \leftarrow cost\_2$;
$\quad\quad\quad\quad \kappa_{temp} \leftarrow \kappa_{temp} + \Delta\kappa$;
$\quad\quad\quad\quad cost\_2 \leftarrow \text{cost}(\kappa_{temp})$;
$\quad\quad\quad\quad$ if $cost\_2 < cost\_1$ then
$\quad\quad\quad\quad\quad\quad \kappa_{rel} \leftarrow \kappa_{temp}$;
$\quad\quad\quad\quad\quad\quad continue \leftarrow$ true;
$\quad\quad continue \leftarrow$ true;
$\quad\quad \kappa_{temp} \leftarrow \kappa_{temp} - 2\Delta\kappa$;
$\quad\quad cost\_2 \leftarrow \text{cost}(\kappa_{temp})$;
$\quad\quad \Delta\kappa \leftarrow \Delta\kappa/2$;
return κ_{rel}

Figure 1. *Arc adaptation algorithm.* One of the typical GESTALT arcs is chosen as an initial estimate for the optimization procedure represented in the algorithm. The approach represents a variant of sampled subgradient descent that can be quite effective in representations commonly encountered in applications such as obstacle avoidance and navigation. The algorithm is invoked in a loop that iterates over all arcs as part of standard arc evaluation in GESTALT.
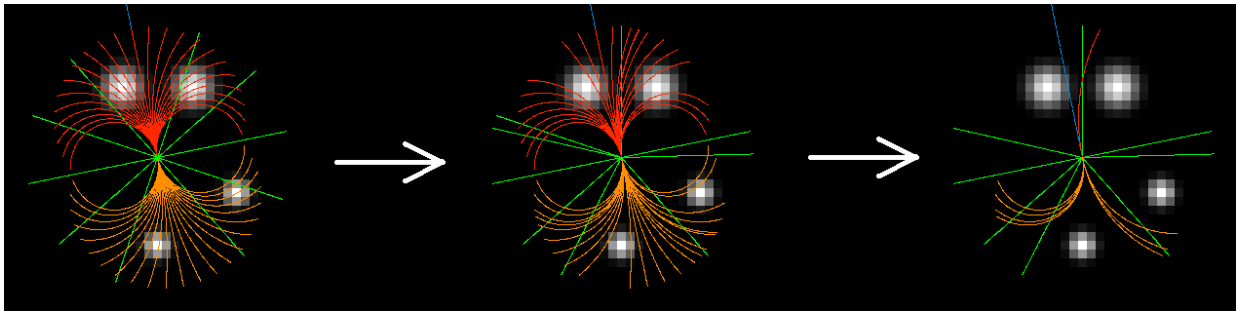
Figure 2. *Arc adaptation results.* The figure demonstrates the application of the algorithm in Figure 1 to a standard GESTALT set of arcs. Forward arcs are represented in red, backward in orange, and point turns in green; thin blue line is the rover's heading. Cost of traversal (real-valued) is represented in grayscale intensity, where white denotes highest cost. On the left, the original arc set is shown. The middle of the figure shows an intermediate stage of adaptation. On the right, the arcs adapted to the cost minima are shown. The case of redundant arcs converging to the same arc is easy to detect by simply evaluating arc curvature; arcs with similar curvatures (based on an intuitive threshold) are removed.
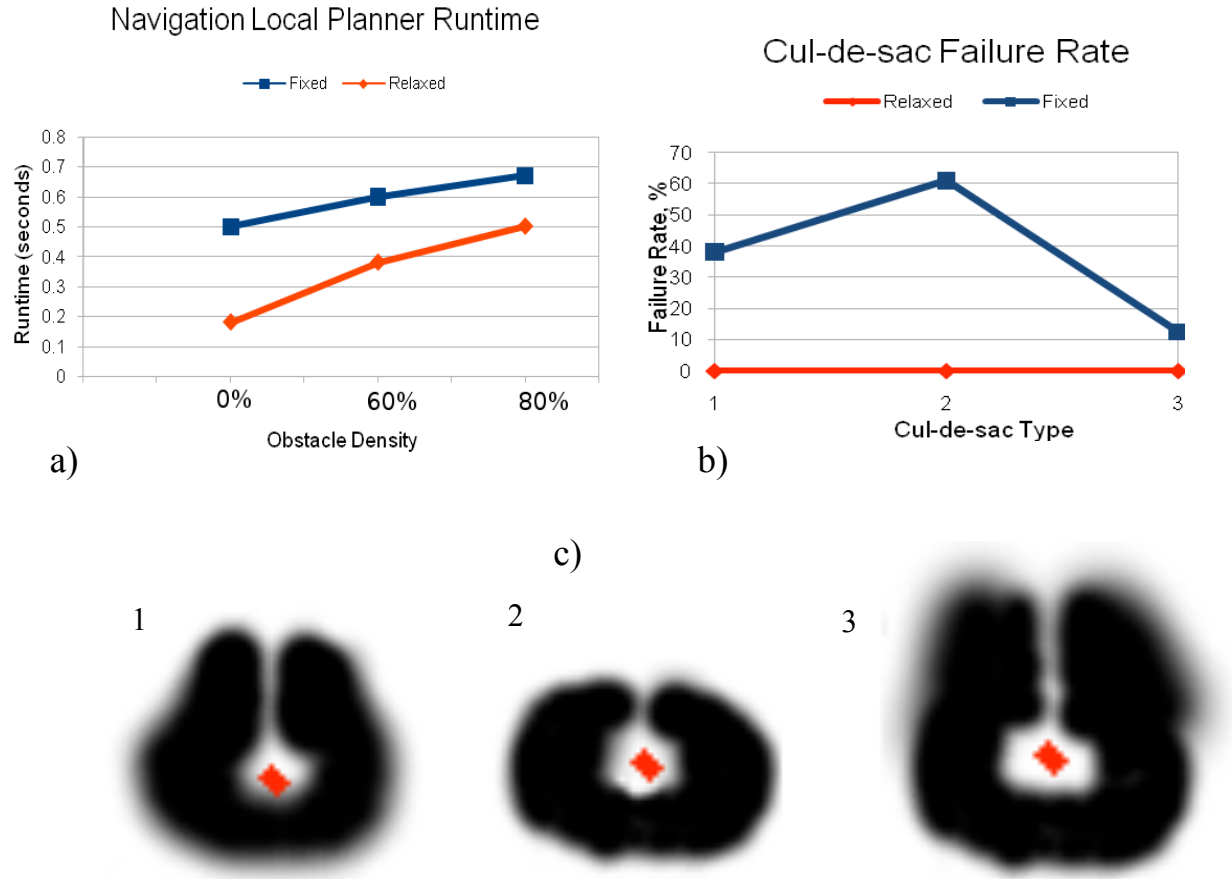
Figure 3. *Arc adaptation results.* The subfigure a) represents the results of a comparative study of simulated rover navigation. Two types of navigation are compared: standard set of GESTALT arcs (consisting of 23 fixed arcs; blue line) and a set of only 7 arcs, endowed with adaptation capacity. Despite the additional computation overhead of adaptation, by virtue of considering only 7 arcs the adaptation-enabled navigation consistently requires less computation. The subfigure b) illustrates a set of experiments where the rover was placed inside one of the three cul-de-sacs (subfigure c). Navigation goal was placed outside of the cul-de-sacs, and the rover was required to maneuver out. Similar to a), the blue line represents the standard GESTALT arc set, while the red line represents a set of 7 arcs endowed with adaptation capacity. The latter was able to guide the robot out of all cul-de-sacs, while the former was not able to compute the required motion in a significant percentage of trials (10%–60%). The details of the cul-de-sacs are shown in subfigure c). The red rectangle is the comparative size of the rover. The grayscale represents the height of the rocks around the rover: black is untraversable, while shades of gray are still within traversability capacity. The exits out of cul-de-sacs were chosen to be narrow enough such that the rover would fit at close tolerance.

ATHLETE LUNAR PROTOTYPE ROVER MOBILITY CHARACTERIZATION AND CONTROL DEVELOPMENT

Jaret B. Matthews (PI), Robotic Hardware Systems Section (347)
Brian Wilcox (Co-PI), Autonomous Systems Architecture and
Program Development Section (340)
David S. Wettergreen (Co-PI), Robotics Institute, Carnegie Mellon University
Scott J. Moreland, Mechanical Engineering Department, Carnegie Mellon University

A.   OBJECTIVES

The objective was to develop a test program for the JPL ATHLETE mobility system that investigated, specifically, rover/terrain interaction with the goal of improving control over extreme terrain and optimizing efficiencies for average terrain. In addition, vehicle mobility characterization was to be completed. This work would forward the understanding of the prototype rover through laboratory and field testing in lunar relevant terrain. Quantitative values were considered essential to the study, as the ATHLETE mobility platform is a relatively new system compared to the well studied rocker-bogie system implemented on all JPL flown rovers.

The second half of the proposed work was to focus on changes to the ATHLETE rover low-level locomotion control. The methods of improving the current control techniques were to arise from the field testing and experimentation. These developments will aid in reaching the goal of increased terrain types the rover can traverse in order to reach areas of scientific interest.

B.   APPROACH AND RESULTS

1.   Test Preparation and Experimentation

Application of mobility performance metrics required the use of experimental apparatus to control the vehicle loading and traction situation. Measurement of the vehicle response to applied loading and terrain is used to quantify performance. The apparatus to conduct drawbar pull and slope ascent experiments were designed and fabricated or reconfigured COTS hardware was used.

a.   *Design/fabrication of portable variable resistance sled with load sensing.* Drawbar pull tests require an applied load in the opposing direction of the vehicles motion. A sled design was chosen for its ability to be easily used in either the Mars Yard or field test locations. This sled was able to apply a load up to 50% of the vehicle's weight (15000N). A pulley system was utilized to apply a constant equal load to all six or all four wheels at a point of

interaction close to the ground. A load cell interfaced to a microcontroller with wireless capability was implemented in order to simultaneously log resistance load and vehicle slip on a laptop located next to the tracking station.

b.   *Development of location/slip sensing.* The vehicle position had to be constantly tracked during tests in order to make wheel slip measurements. This required data to be collected at ~1 Hz with a position accuracy of <3 mm. A Leica TCRA 1103 Total Station (3 arcseconds angular accuracy, 1 mm distance accuracy) was used to track a 3-D prism on board the ATHLETE prototype. Software was developed to control the total station via a laptop and log load data wirelessly from the resistance sled.

c.   *Configure terrain for drawbar pull tests.* Loose soil/sand of low bearing strength is required for the mobility experimentation. As all test results were to be compared to only others conducted during this test program, exact soil properties were not required. A relative low shear strength and low bearing strength were all that were required to produce adequate wheel slippage and sinkage relative to planetary rovers. As such, the sand located in the JPL Mars Yard would suffice once fully loosened. A skid loader was used to loosen (turn over) a patch of sand (4 m × 23 m × 0.30 m deep).

d.   *Develop technique for preparing soil to relevant conditions.* A repeatable method of achieving soil properties was tested. This consisted of completely loosening the material, raking level, and then compacting with a heavy flat surface. A cone penetrometer was used to assure repeatability of the compaction resistance curve for each experimental run.

e.   *Drawbar Pull Experimentation.* Vehicle difficulties encountered during initial drawbar pull experiments with the ATHLETE prototype rover (SDM-B) resulted in the delay of testing in the Mars Yard. The team opted to use the newly developed TriATHLETE prototype as these tests are relatively high stress for the previous mobility system. This new prototype has much larger actuators, and power system and structure better suited to these experiments. The demonstration of the TriATHLETE prototype at an analogue field site provided the best opportunity to complete the drawbar pull tests but poor weather prevented from completion of these mobility characterization experiments.

2.   Control Development

As no data were collected on the locomotion system performance and difficulties encountered using the vehicle prototypes, there were no changes or development of the mobility system control.

C.   SIGNIFICANCE OF RESULTS

Performing the planned test program can aid in forwarding understanding of the ATHLETE mobility system and how it interacts with lunar-like terrain. Applying this knowledge will lead to more efficient control methods that can permit the traverse of extreme terrain, under lower risk. The characterization of the ATHLETE platform will also aid in planning as a result of

knowledge of terrain capabilities and rover requirements such as accurate power values or vehicle limitations for various configurations.

The development of vehicle test apparatus and procedure now allows for ongoing evaluation of the ATHLETE lunar rover and follow-on research. This can be undertaken in the Mars Yard or at annual analog site field trials. Furthermore, the ability to utilize the mobility system as a platform for general vehicle experimentation is now possible. The high d.o.f chassis can be reconfigured to mimic many mobility platform types and comparative studies can now be conducted with ease in order to analyze various locomotion modes (walking, rolling, hybrid braked/roll).

In addition to technical gains, the proposed work has resulted in strengthened ties between Carnegie Mellon University's Robotics Institute and JPL's Robotic Hardware Systems group. The SURP award has also provided an avenue for valuable educational and practical experience for students directly involved with this effort.

D.   FINANCIAL STATUS

The total funding for this task was $25,000, of which $23,535 has been expended.

E.   ACKNOWLEDGEMENTS

Thank you to Julie Townsend, David Mittman, Jack Dunkle, Matthew Heverly, and Ross Finman for their aid in many aspects of this work.

F.   PUBLICATIONS

None.

G.   REFERENCES

None.

Figure 1. Drawbar pull experimental set-up in Mars Yard with ATHLETE Lunar Prototype.

DARTMOUTH GREENCUBE 2

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1308

Anthony J. Mannucci (PI), Tracking Systems and Applications Section (335)
Kristina A. Lynch (Co-PI), Department of Physics and Astronomy, Dartmouth College

A.   OBJECTIVES

The GreenCube project stems from Professor Kristina Lynch's interest in small, autonomous science payloads for multipayload auroral sounding rockets, and from Professor Robyn Millan's interest in small, CubeSat-like orbiters for future science missions.

The goals of the GreenCube project are (1) to maintain a scientifically interesting, student-driven, balloon-borne CubeSat program in Dartmouth Physics; (2) to incorporate new design features into small payloads for low-cost access to space (LCAS)–class auroral sounding rocket proposals by Prof. Lynch; and (3) on a longer timescale, to incorporate designs into Professor Millan's future plans for small orbiters, which potentially could include student-driven CubeSats.

The specific goal of GreenCube 2 was to use the GreenCube infrastructure to collect and understand data, especially GPS and temperature data, pertaining to gravity waves. GreenCube 2 was the first GreenCube mission to pursue a true science goal rather than just infrastructure development.

B.   APPROACH AND RESULTS

GreenCube 2, described here, is a follow-on to the GreenCube 1 project. The students found that preliminary measurements from the GreenCube 1 flight GPS system showed indications of gravity waves in the wind velocity measurements; determined that a higher data rate was necessary and increased it from a 30-second cadence to a 6-second cadence; revisited the temperature measurements from GreenCube 1 and changed the circuitry to support exterior thermistor measurements; built a second copy of the GreenCube; and launched, flew, and recovered both payloads in August of 2009.

Two adjacent GreenCubes were flown on two high-altitude sounding balloons that reached approximately 90,000 feet altitude before bursting. The payloads then descended via parachute and were retrieved using the real-time GPS track received through the ham radio system. The balloons flew over the Presidential Range of mountains and were recovered in Maine. The flight time was approximately 2 hours.

GPS data were transmitted to the ground crews in real time over the course of the flight. GPS position vectors and timestamps were recorded and transmitted every 5 seconds, as shown in Figure 1. Much like GreenCube 0 and GreenCube 1, the payloads of GreenCube 2 ascended in a roughly linear fashion. After the balloon burst at the flight apogee, the balloons descended in a roughly exponential profile. As the atmosphere became denser at lower altitude, the parachute created more drag, lowering the ascent rate. Both balloons burst between 25 and 28 km in altitude, well within the preflight prediction of 80,000 to 90,000 feet.

The velocity of the balloon was derived from the recorded position and time data by dividing the distance between GPS position coordinates by the time delay between them (usually 5 seconds). In this manner, both the speed of the payload and its heading were calculated. Horizontal velocity profiles are shown in Figure 2. The team made an important assumption that the velocity of the payload at any given moment is the same as the wind speed, i.e., the payload accelerates nearly instantaneously with the wind. The payload profiles look almost identical, indicating that they ascended through similar atmospheric features. As the balloons approach 5 km in altitude, they enter the jet stream and accelerate very rapidly. After achieving about 15 km in altitude, the balloons begin to slow down. Gravity waves may be causing the many small-scale fluctuations seen in these horizontal velocity profiles.

Also flown aboard the payloads were five thermistors that recorded temperature data. Data from the thermistors were transmitted approximately every 10 seconds. Before analyzing the temperature data, the data from all five thermistors were averaged so that one temperature profile was derived from each payload. The data were also converted into potential temperature to follow convention (Figure 3). Potential temperature θ is defined as

$$\theta = T \left(\frac{P_0}{P} \right)^{\frac{R}{c_p}},$$

where P is the pressure, P_0 is the reference pressure, R is the gas constant of air, and c_p is the specific heat at constant pressure. The reference pressure used was 1000 millibars. Because the GPS data and temperature data were handled by two different systems, they therefore were reported at different rates. The time stamps sent out on each radio transmission represented times corresponding to the GPS data but not necessarily the temperature data. The temperature data were splined with the GPS data to determine when and at what altitude the temperature data were recorded.

To start our gravity-wave analysis, we transformed our velocity data into components. A look at the compass heading of the balloon over the course of its ascent shows that the balloon oscillated in direction around a 120-degree heading. Therefore, we designed a new coordinate system that better reflected the direction of the balloon, in which the "along" velocity shows movement in the direction of the prevailing winds, and the "across" velocity registers movement perpendicular to the prevailing winds. When the new coordinate system is applied, the oscillations are more pronounced and the heading changes are more apparent. The fluctuations in horizontal velocity are strongest perpendicular to the balloon's path.

Combining the position and velocity data on a quiver graph allows for a glimpse into two slices of the atmospheric velocity vector field above Mount Washington (Figure 4a). At first

glance, the velocity vector profiles look very similar. Features are the same across lines of equal altitude. There is no obvious time or range dependence. However, a closer look shows some slight changes along lines of equal altitude.

For example, along the 14-km altitude contour, a peak in the cross-component velocity can clearly be seen just above the contour on both payloads as they ascend through this area. However, when payload 1 descends, the peak occurs exactly on the contour, and payload 2's peak occurs below this contour. The change in altitude of this atmospheric feature is indicative of either a change with respect to time or distance of this feature.

We can use this phase change to measure the horizontal wavelength of this feature. The change in altitude of two vertically propagating wave structures can be used to calculate the horizontal wavelength of the structure, as shown in Figure 4b. Because the horizontal wavelength is proportional to the change in altitude, the peak-to-peak change in altitude of one full vertical wavelength can be used to calculate the horizontal wavelength: $\lambda = \frac{\Delta altitude}{slope\ of\ red\ line}$.

In Figure 4a, we plotted lines connecting similar features. Rather than connecting peaks, however, we chose to connect the nodes together. The nodes represent a change in the across velocity from southeast to northwest. The lines with the greatest slope are located between 13 to 16 km in altitude. These lines describe the atmospheric features with the greatest change. The phase change becomes zero at increased altitudes. The line around 14.5 km had the greatest slope at −17.36 m per km distance. The peak-to-peak vertical wavelength of payload 1's ascent measured 1.3 km. Therefore, the minimum horizontal wavelength we observed was 76.5 km. The balloons themselves only collected reliable data above 10 km altitude and less than 40 km. Therefore, the structure is so large that horizontal distance is a negligible factor in determining the across velocity.

Starting at 15 km altitude, the balloon makes repeated fan-like patterns in which the balloon sharply changed direction and then gradually returned to its normal heading. The largest of such fluctuations occurs at 15 km, corresponding with the tropopause and therefore the largest change in temperature. The other prominent sharp accelerations occur at 17.8 and 20.5 km altitude, which are the same regions that correspond to the second and third largest changes in temperature. This could be indicative of an atmospheric shear layer — an area of the atmosphere where the velocity of the wind is vastly different from the layers above and below it. This sudden change in velocity of the gas would change the balloon's heading and cause it to record a very different temperature. To conclude, the large changes in temperature and velocity occur together, and indicate that the payload is passing through a discontinuous shear layer.

In addition to the GreenCube payloads, each balloon also carried a commercial camcorder on its lower secondary payload (designed to carry an emergency locator transmitter (ELT).) These cameras captured HD video of the Earth from the balloon altitudes, including images of cloud formations that bore the signs of atmospheric gravity waves (Figure 5).

## C.	SIGNIFICANCE OF RESULTS

The GreenCube team met its science goals by collecting temperature and GPS data as well as camera footage, which were then analyzed in an attempt to locate gravity-wave signatures.

The GreenCube program addresses JPL's interests in enhancing student preparation for a professional career in space systems/science at JPL or elsewhere. The students have gained experience with instruments of particular interest to JPL, such as GPS and magnetometers, as well as with analysis techniques for multipoint in situ geophysical observations.

## D.	FINANCIAL STATUS

The total funding for this task was $25,000, all of which has been expended.

## E.	PUBLICATIONS AND PRESENTATIONS

[A]	Umair Siddiqui, Max Fagin, Sean Currey, and Amanda Slagle, "Dartmouth GreenCube: Multiple Balloon Measurements of Gravity Waves over New Hampshire," Colloquium, Dartmouth College, Hanover, New Hampshire, November 20, 2009.

[B]	Phil Bracikowski, Max Fagin, Sean Currey, Umair Siddiqui, Kristina Lynch, Amanda Slagle, "GreenCube and RocketCube: Low-Resource Sensorcraft for Atmospheric and Ionospheric Science," AGU Fall Meeting, San Francisco, California, SM33C-1586, December 2009.

[C]	Phil Bracikowski and Umair Siddiqui, "GreenCube: A Student Multiple Small Payload Project — A First Step," CalPoly CubeSat conference, San Luis Obispo, California, spring 2009.

## F.	REFERENCES
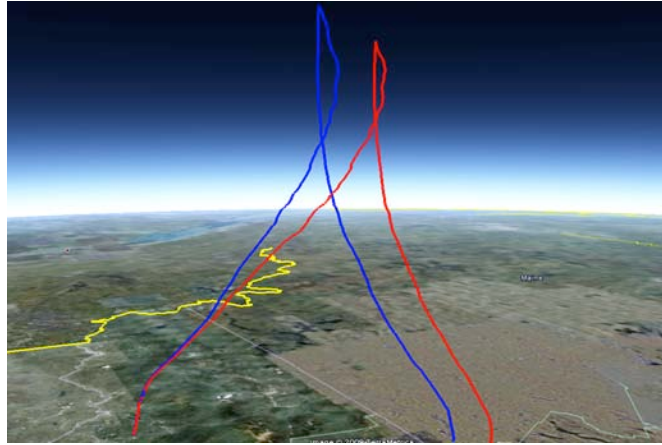
None.

G.   FIGURES

Figure 1. Plot of all position coordinates for both payloads created using Google Earth. The blue trajectory belongs to the first balloon launched (payload 1) and the red belongs to the second balloon launched (payload 2). Payload 2 was launched 90 seconds after payload 1.
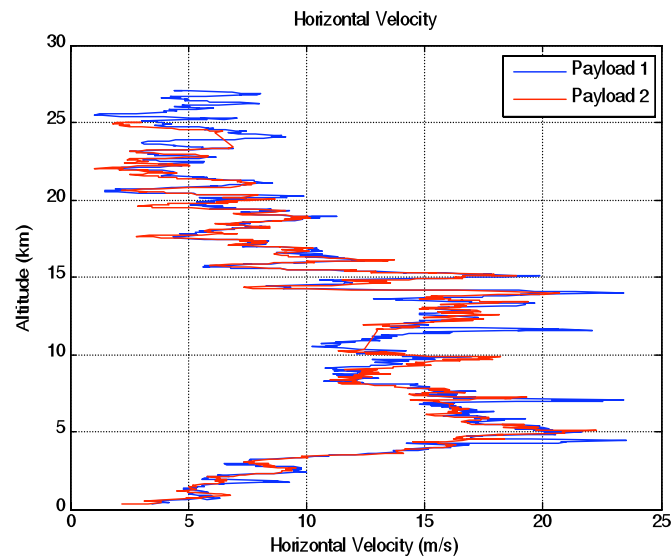

Figure 2. This graph shows the horizontal velocity profiles of both payloads as they ascended to apogee.
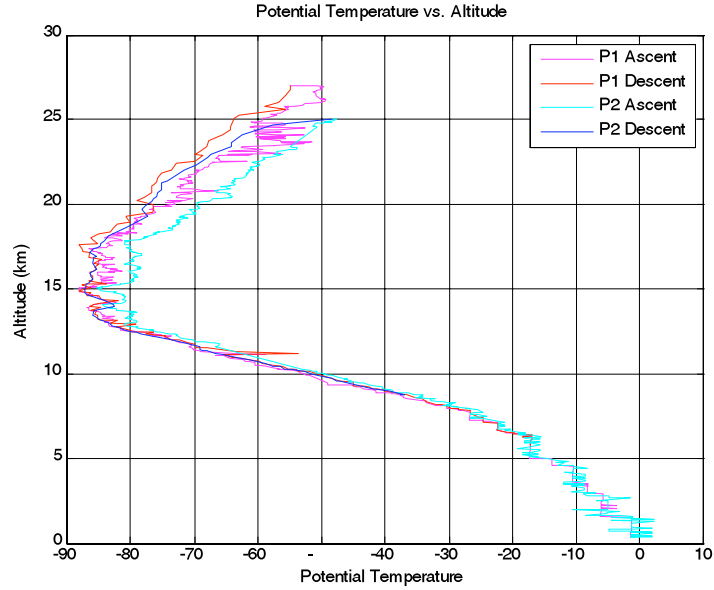
Figure 3. Average potential temperature for each payload plotted against altitude.
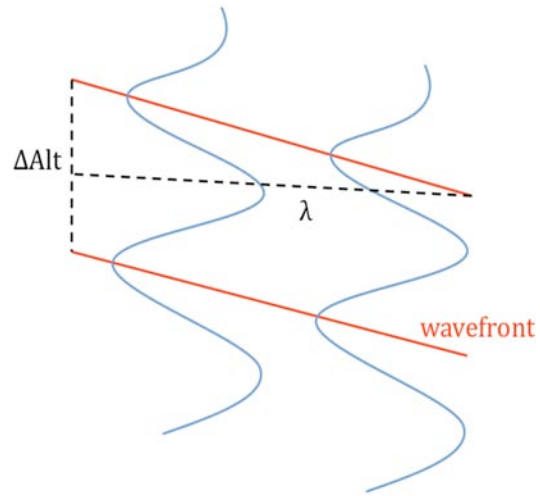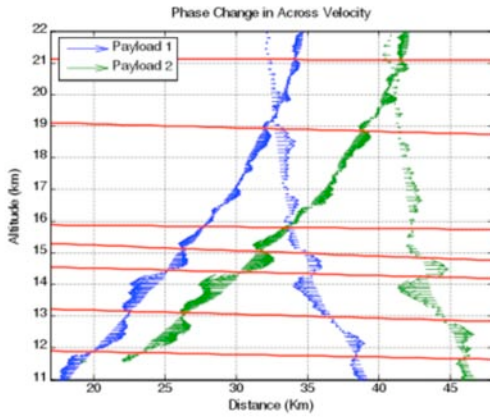

Figure 4. (a) A quiver plot showing the "across" (perpendicular to prevailing winds) component of payload velocity plotted along payload trajectory. The red lines connect atmospheric features detected by payload 1 with corresponding features detected by payload 2. Note that the lines are not perfectly horizontal, suggesting that the two payloads encountered the same features at slightly different altitudes; (b) shows a clearer example of the concept. In this figure, the vertical wavelength λ is given by $\lambda = \frac{\Delta altitude}{slope\ of\ red\ line}$.

Figure 5. Gravity wave features captured by the cameras.

SPACE SYSTEMS PRODUCT DEVELOPMENT: EDUCATING THE NEXT GENERATION OF SPACE SYSTEMS ENGINEERS

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1310

Simon Nolet (PI), Guidance, Navigation and Control Section (343)
David W. Miller (Co-PI), Department of Aeronautics and Astronautics,
Massachusetts Institute of Technology

A.   OBJECTIVES

The objective of the Massachusetts Institute of Technology (MIT) satellite design class was to immerse juniors and seniors in the Department of Aeronautics and Astronautics in the full lifecycle of developing an aerospace project. A multi-semester course was designed to provide future engineers, and potential JPL employees, with an understanding of the impact of requirements definition and design decisions on the performance, manufacture, integration, test, and operations of the system. This course enables the development of problem solving skill-set by active hands-on application to a complex design problem, where the students are given the chance to gain experience with real hardware. Students are guided through focused analysis and application of the design process (conception, design, implementation and operation phases). They are given personal, interpersonal, team and organizational leadership exposure through interaction with team members, communications via different mediums, and programmatic challenges. The SURP/DRDF funds for this award were used to support the course during the Fall 2009 semester.

B.   APPROACH AND RESULTS

1.   The CASTOR Satellite Program

The Cathode/Anode Satellite Thruster for Orbital Repositioning (CASTOR) satellite project was used to achieve the objectives outlined in Section A. CASTOR is a small satellite designed to achieve 1,500 hours of on-orbit operation using an electric propulsion system called the Diverging Cusped Field Thruster (DCFT). This thruster was developed at MIT's Space Propulsion Laboratory and utilizes three permanent magnets of alternating polarities to accelerate xenon ions [1]. The DCFT's diverging structure and cusped magnetic field design is intended to lower the degradation rate for low-powered plasma thrusters. Additionally, CASTOR is an entrant in the University Nanosatellite Program (UNP), in which 11 U.S. universities are competing for a launch opportunity. In August 2009, the CASTOR program completed its preliminary design review (PDR). Much of the fall 2009 semester was dedicated to incorporating PDR feedback and to developing the CASTOR design in preparation for the program's critical design review (CDR) in spring 2010.

2. Program Structure

The CASTOR program was divided into eight subsystems, each focusing on a unique aspect of satellite design. This structure is outlined in the organizational chart shown in Figure 1. At the top, the systems team coordinates with the various subsystem teams to ensure the program produces an integrated satellite design. The avionics and power teams are responsible for interfacing with each electronic component on the satellite while ensuring the satellite has the necessary computing and power capacities needed to function on orbit. The communications team ensures the satellite can transmit and receive data as it passes over ground stations, and the propulsion team oversees the test and integration of the DCFT. The Attitude Determination and Control System, and Guidance, Navigation, and Control (ADCS/GNC) team designs the control systems for the satellite, and the Science and Payload team develops on-orbit monitoring and test procedures for the thruster. Lastly, the structures and thermal teams work to build a mass-efficient, rigid structure that provides attachment points for each component while maintaining components within their operational temperature ranges.

3. Goals and Achievements

The CASTOR program set and accomplished a series of challenging goals for the fall 2009 semester. Other than the mandatory reviews [A–D], notable achievements included a major redesign of the satellite's structure, selecting the satellite's processor and operating system, resolving battery charge issues, and prototyping torque coils and sun sensors. Figures 2 and 3 show a rendering of the latest CASTOR design, in both stowed and deployed configurations. All goals were accomplished with the exception of a thruster test fire and characterization campaign. Testing was delayed because the test chamber was occupied by another propulsion system. This testing does not currently impose any programmatic delays and will be completed during spring 2010.

Throughout the fall 2009 semester, the CASTOR team learned several important lessons. Two examples are provided as an illustration. In response to PDR feedback, the structures team expanded its design focus to better accommodate manufacturability and assembly. Thanks to these efforts, the team was able to fabricate, assemble, and test a structural engineering model in less than one month. Figures 4 and 5 show pictures of the structural engineering model vibration testing. Next, the avionics team learned a great deal about processor sizing for mission requirements. The team discovered that an F-12 single-board computer was not required to fulfill the mission's processing needs. Instead, this component was replaced with three microcontrollers saving 3 W of power and 0.5 kg of mass.

C. SIGNIFICANCE OF RESULTS

For years, industry leaders have bemoaned the fact that engineering education has drifted from real-world demands on engineers. The MIT Department of Aeronautics and Astronautics has created the landmark Conceive, Design, Implement, and Operate (CDIO) curriculum to graduate students possessing the specific skills and experiences that industry has identified as vital for new engineers. JPL's partnership in this effort ensures that MIT can continue preparing

students to enter the workplace uniquely ready to conceive, design, implement, and operate complex product systems. This course is a unique, hands-on, design-build class where undergraduates use the full set of CDIO skills they have acquired by applying them to the complete development cycle of an aerospace vehicle. Therefore, students of the class (and potential JPL hires) are better prepared to face the challenges that aerospace industries face daily in the real world.

D.   NEW TECHNOLOGY

DRDF funding supported the development of the MIT CASTOR Satellite Program, whose mission is to validate the performance and application of DCFT technology developed at MIT's Space Propulsion Lab under a different sponsorship [1]. This will be achieved by taking on-orbit state data to compare the degradation experienced by the DCFT during orbital maneuvers to that of similar technologies.

E.   FINANCIAL STATUS

The total funding for this task was $25,000, of which $23,200 has been expended.

F.   ACKNOWLEDGEMENTS

The CASTOR Program wishes to thank Col. (ret.) John Keesee, Dr. Alvar Saenz-Otero, and Mr. Paul Bauer for their mentorship and technical expertise in designing the CASTOR satellite.

G.   PUBLICATIONS AND PRESENTATIONS

[A]   CASTOR Preliminary Design Review, presented at Logan, Utah, 14 August 2009.

[B]   CASTOR Initial Status Review, presented at Cambridge, Massachusetts, 8 October 2009.

[C]   CASTOR Midterm Status Review, presented at Cambridge, Massachusetts, 12 November 2009.

[D]   CASTOR Final Status Review, presented at Cambridge, Massachusetts, 3 December 2009.

H.   REFERENCES

[1]   Daniel George Courtney, "Development and Characterization of a Diverging Cusped Field Thruster and a Lanthanum Hexaboride Hollow Cathode," Master's Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, June 2008.
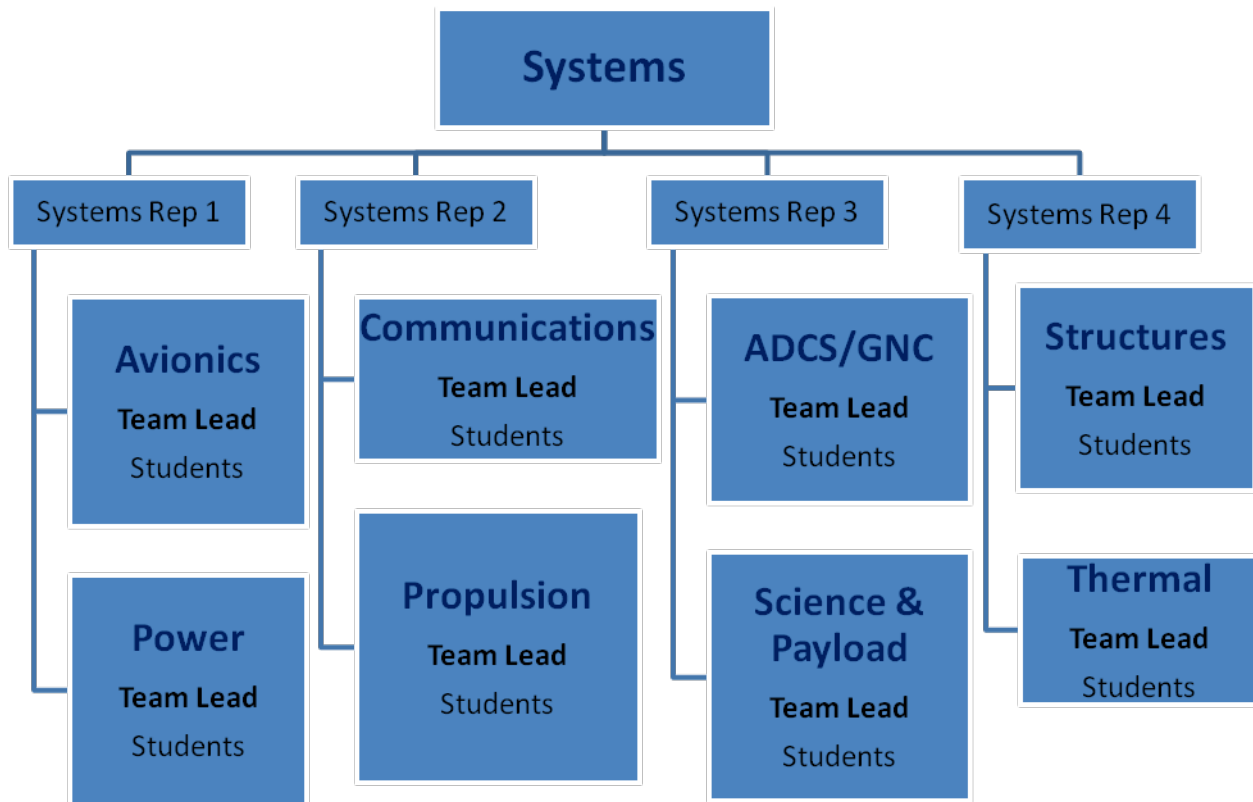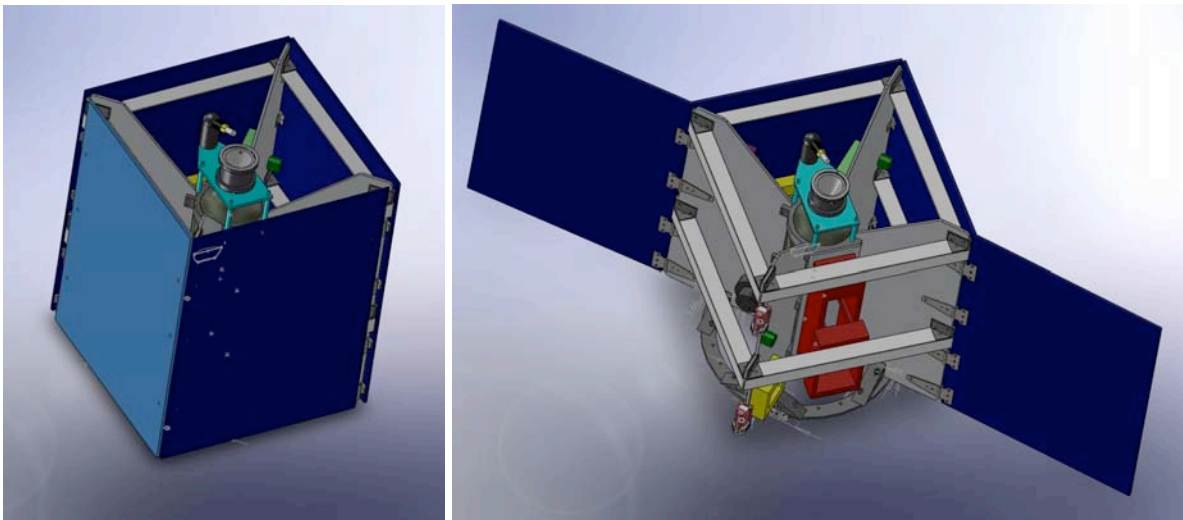
I. FIGURES

Figure 1. CASTOR subsystems organization.

Figures 2–3. CASTOR rendering in stowed (left) and deployed (right) configurations.
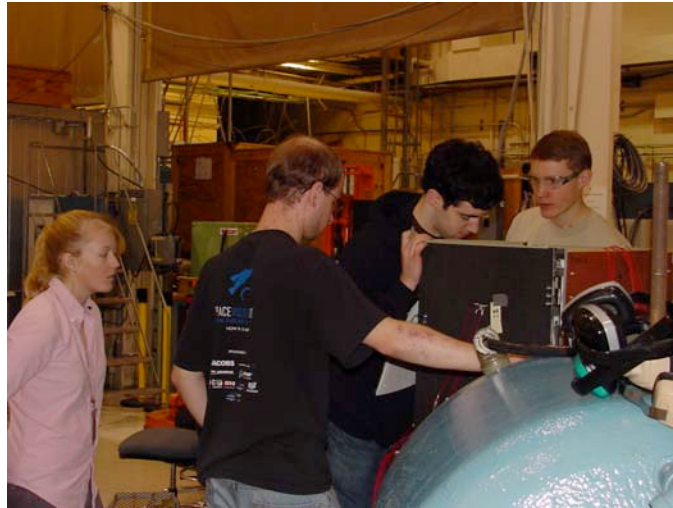
Figure 4. Students preparing the CASTOR structural engineering model for vibrational testing.
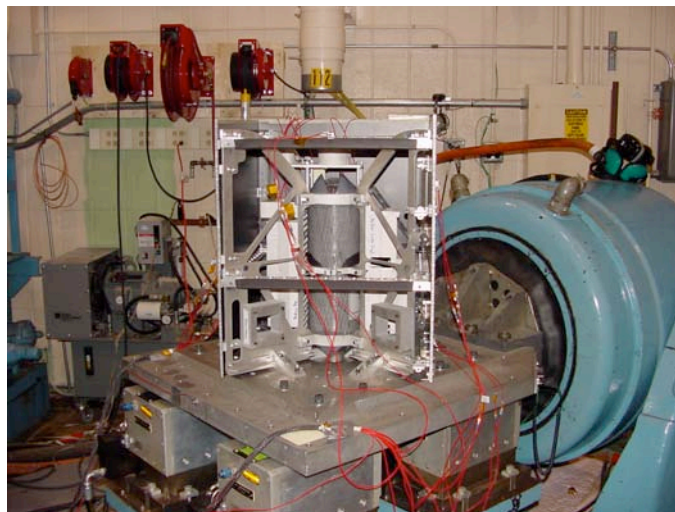

Figure 5. The CASTOR structural engineering model at vibration testing.

A MULTI-SENSOR PERSPECTIVE ON UPPER TROPOSPHERIC HUMIDITY

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1317

Annmarie Eldering (PI), Earth Atmospheric Science Section (328)
Brian H. Kahn (Co-PI), Climate, Oceans, and Solid Earth Section (324)
Frederick W. Irion (Co-PI), Climate, Oceans, and Solid Earth Section (324)
William G. Read (Co-PI), Earth Atmospheric Science Section (328)
Kuo-Nan Kiou (Co-PI), University of California at Los Angeles/Joint Institute for
Regional Earth System Science and Engineering

A.   OBJECTIVES

The scientific objective of this research was to quantify the role that upper-tropospheric (UT) tropical clouds play in controlling the upper tropospheric humidity (UTH) structure using the A-Train satellite measurements of water vapor, clouds, and temperature. Specifically, we will identify moistening and drying processes and determine the cloud regimes that correspond to these processes.

The objectives were broken into two tasks: 1. Create joint AIRS, MLS, CloudSat, and CALIPSO dataset of collocated humidity and cloud retrievals. Joining the datasets will take part in three stages, combining: i) AIRS and MLS, ii) AIRS, CALIPSO, and CloudSat, and 3) all four instruments. 2. Trajectory analysis (forward and backward) to better understand sources and sinks of UTH and the evolution of UT clouds.

B.   APPROACH AND RESULTS

1.   Unified Dataset

We proposed to create a unified dataset of thermodynamical variables, i.e., temperature, water vapor (H_2O), and relative humidity (RH) from soundings retrieved from the Atmospheric Infrared Sounder (AIRS) and the Microwave Limb Sounder (MLS), through analyzing each instruments averaging kernels (AKs). We find that AKs do have skill in identifying the regions in the atmosphere where AIRS has sensitivity to H_2O. We analyzed the level of the peak of the averaging kernel: kernels peak around the correct retrieval level up to ~261 hPa, an indicator that AIRS has sensitivity at the correct regions of the atmosphere. The total verticality (TV) and the local verticality (LV) were analyzed. From these results, we found that AIRS starts to drastically lose sensitivity to H_2O above 261 hPa (in altitude). This is consistent with previous studies [1,2] that showed AIRS and MLS seem to agree best around 261 hPa.

With this information, we were able to splice the AIRS and MLS H_2O profiles using each instrument's LV as coefficients to compute a weighted mean between the two soundings. All the profiles have smooth transitions between all pressure levels and contain no unphysical "kinks." The joint distributions (Figure 1) can be interpreted as the H_2O lapse rate (as a function of pressure) reported as a percentage change from the H_2O value at the lower adjacent pressure level; we call this DQ. We see that for the MLS and AIRS data alone there are no discontinuities in DQ. The same is true for the newly combined dataset shown in panel (c). This gives us confidence that we spliced the data together in a manner that does not degrade the underlying physics in the profiles. Furthermore, the new H_2O values at the spliced levels (between 215 hPa and to the top of atmosphere) do not fall outside the specified accuracies for each instrument (not shown). This is not surprising as the MLS has been shown to be slightly dry [2] at these pressure levels, while the AIRS has been shown to be slightly moist at these levels [1]. Our splicing technique basically constrains the H_2O values to lie between the MLS and AIRS retrievals. From this, we are able to construct a climatology of temperature and humidity that spans the entire atmospheric column.

We have also collocated all the available CloudSat and Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) cloud profiles from June 2006–February 2009. The CALIPSO data beyond February 2009 is not available for scientific analysis. In section B.3 (Radiative Impact of Clouds), we explain in more detail the usage of the cloud and thermodynamic profiles to compute heating rates in the upper troposphere/lower stratosphere (UTLS).

2.   Analysis of Unified Dataset

We originally proposed to quantify the joint distributions of cloud and thermo-dynamical properties of the tropical UTLS. Although we have some results along these lines, we have first focused on quantifying the interannual variability of temperature and humidity in the tropical atmosphere. Specifically, we are quantifying the effects of QBO and the ENSO on UTLS humidity. Figure 2 shows the time series for H_2O, temperature, and RH interannual anomalies from the surface (1013 hPa) to 10 hPa, for the 5S–5N latitude band globally. We find that the A-Train datasets do clearly identify the ENSO and QBO (see Ocean Niño and QBO indices in Figure 2d). The time series is consistent with the ENSO being the dominant interannual mode from the surface and throughout the free-troposphere, and the QBO dominating the middle to lower stratosphere. However, in order to isolate each mode, we compute the same time series between 5S–5N but in the Tropical Western Pacific (TWP, 120E–170E) and Nino3.4 region (120W–170W), since ENSO has a particularly strong impact on these regions in the tropics. Since the QBO is zonally symmetric [3], we can subtract one of the local time series from the global time series in Figure 2. Figure 3 shows the results of that calculation in the Nino3.4 region. All the panels follow Figure 2 but represent the difference between localized and global anomalies. Figure 3, panels (b) and (c), we see a clear ENSO signal near 100 hPa (around the tropopause). The strongest signal is during the La Niña event between the end of 2007 and 2009, a particularly strong ENSO event (see Figure 3d). Another interesting feature is the positive H_2O anomaly we see collocated with the positive temperature and negative RH anomalies. Above this strong signal in the stratosphere the temperature and H_2O signals is negligible, qualifying that the QBO is indeed zonally symmetric. The peak temperature signals of about ±1.5 K correspond to

about $\pm10\%$ (±0.3 ppmv) change in H_2O; this is $\sim8\%$ of the mean concentration of H_2O around 100 hPa. Therefore, this ENSO effect on temperature leads to a non-negligible signal in H_2O.

3.   Radiative Impact of Clouds

Another element of the proposal was to quantify the radiative impacts of H_2O and cirrus on the tropical atmosphere. A preliminary attempt at this computation is shown in Figure 4. The data are calculated from May 2008–February 2009 since this is the only period for which AIRS, MLS, CloudSat, and CALIPSO are collocated in space and time. The clear-sky heating rates (panel (a)) basically represent the meridional distribution of the level of zero radiative heating (LZRH), which is primarily driven by H_2O emission in the troposphere; this is represented by the solid black "zero" contours. The locations of the LZRH do not change very much in the tropical clear-sky and are consistent with [4] and [5]. Panels (b) and (c) show the impacts of various clouds on the LZRH. It is clear that clouds dramatically alter the radiative balance in the atmosphere. The locations of maximum heating follow the dash-dotted cloud contours. The cumulative effect of clouds on the clear-sky LZRH is shown in (d).

C.   SIGNIFICANCE OF RESULTS

This research is a significant advancement in the integration of A-train measurements to study large-scale atmospheric processes that impact water vapor, temperature, and clouds. We now have an integrated dataset (the first one that we are aware of that combines AIRS, MLS, Cloudsat, and CALIPSO), and have started to analyze the signature of QBO and El Niño on these fields. The graduate student conducting the research has significantly advanced in his PhD research, and has made strong connections to established researchers in the community. His work has been conducted in residence at JPL, and that has benefited JPL through the research that is being completed.

D.   FINANCIAL STATUS

The total funding for this task was $25,000, of which $21,235 has been expended.

E.   PUBLICATIONS and PRESENTATIONS

[A]   Calvin Liang, "A Combined A-Train Perspective on Upper Tropospheric Humidity and Radiative Heating," Upper Tropospheric Lower Stratospheric Physics and Chemistry session, American Geophysical Union Fall Meeting, San Francisco, California, December 14–18, 2009.

[B]   Calvin Liang, "Characterization of A-Train Water Vapor Sensitivity Through Integration of Averaging Kernels and Retrievals," submitted to *IEEE Remote Sensing*.

F.   REFERENCES

[1]   E. J. Fetzer, W. G. Read, D. Waliser, B. H. Kahn, B. Tian, H. Vömel, F. W. Irion, H. Su, A. Eldering, M. de la Torre Juarez, J. Jiang, and V. Dang, "Comparison of Upper Tropospheric Water Vapor Observations from the Microwave Limb Sounder and Atmospheric Infrared Sounder," *J. Geophys. Res*. **113** (2009), D22110. doi:10.1029/2009JD010000

[2]   W. G. Read, A. Lambert, et al., "Aura Microwave Limb Sounder Upper Tropospheric and Lower Stratospheric H_2O and Relative Humidity with Respect to Ice Validation," *J. Geophys. Res*. **112** (2007). doi:10.1029/2007JD008752

[3]   M. P. Baldwin, L. J. Gray, T. J. Dunkerton, K. Hamilton, P. H. Haynes, W. J. Randel, J. R. Holton, M. J. Alexander, I. Hirota, T. Horinouchi, D. B. A. Jones, J. S. Kinnersley, C. Marquardt, K. Sato, and M. Takahashi, "The Quasi-Biennial Oscillation," *Rev. Geophys*. **39** (May 2001): pp. 179–229.

[4]   T. Corti, B. P. Luo, P. Peter, H. Vömel, and Q. Fu, "Mean Radiative Energy Balance and Vertical Mass Fluxes in the Equatorial Upper Troposphere and Lower Stratosphere," *Geophys. Rev. Lett*. **32** (2005), L06802. doi:10.1029/2004GL021889

[5]   A. Gettelman, P. M. d. F. Forster, F. Masatomo, Q. Fu, H. Vömel, L. K. Gohar, C. Johanson, and M. Ammerman, "Radiation Balance of the Tropical Tropopause Layer," *J. Geophys. Res*. **109** (2004). D07103, doi:10.1029/2003JD004190
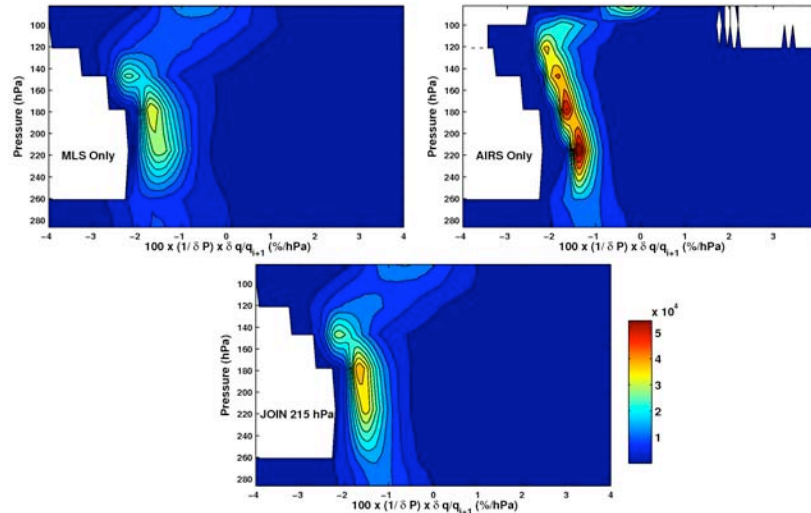
G.   FIGURES

Figure 1. Joint distribution for the value DQ = $1/\delta P * \delta q/q_i + 1$ (multiplied by 100 to get %) for a) AIRS, b) MLS, and c) joined H_2O profiles. The colorbar shows absolute frequency.
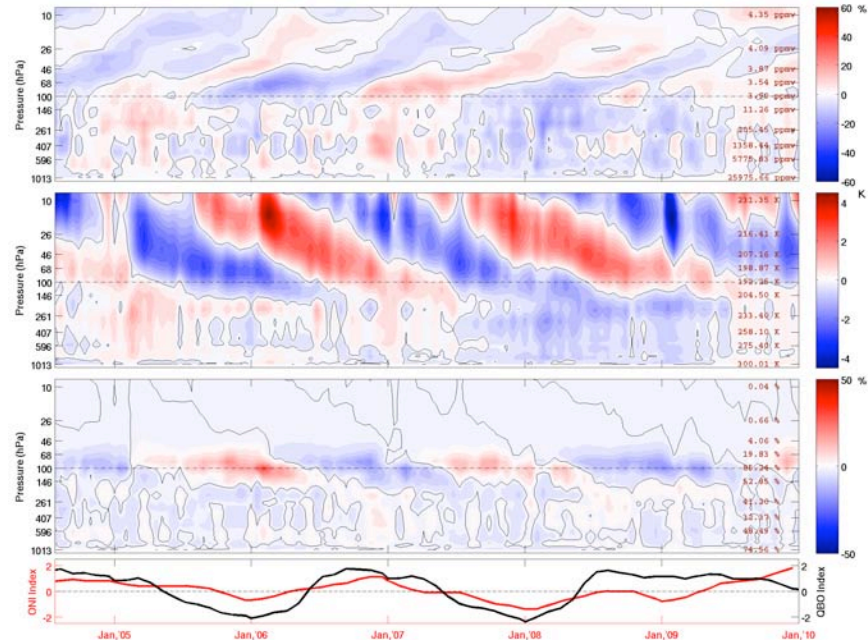
Figure 2. Tropical time series of the interannual variability of: (a) H$_2$O (reported as a % anomaly), (b) temperature anomaly (K), and (c) RH anomaly (%) from the 1013 hPa up to 10 hPa. The red text at the right of each of the contoured panels shows the climatological mean values at the designated pressure levels shown on left ordinate of each panel. Panel (d) shows the Ocean Nino Index over the Nino3.4 region and the QBO index at 50 hPa over Indonesia. The anomalies are zonal means from 5S–5N. Black lines are the "zero" contours.
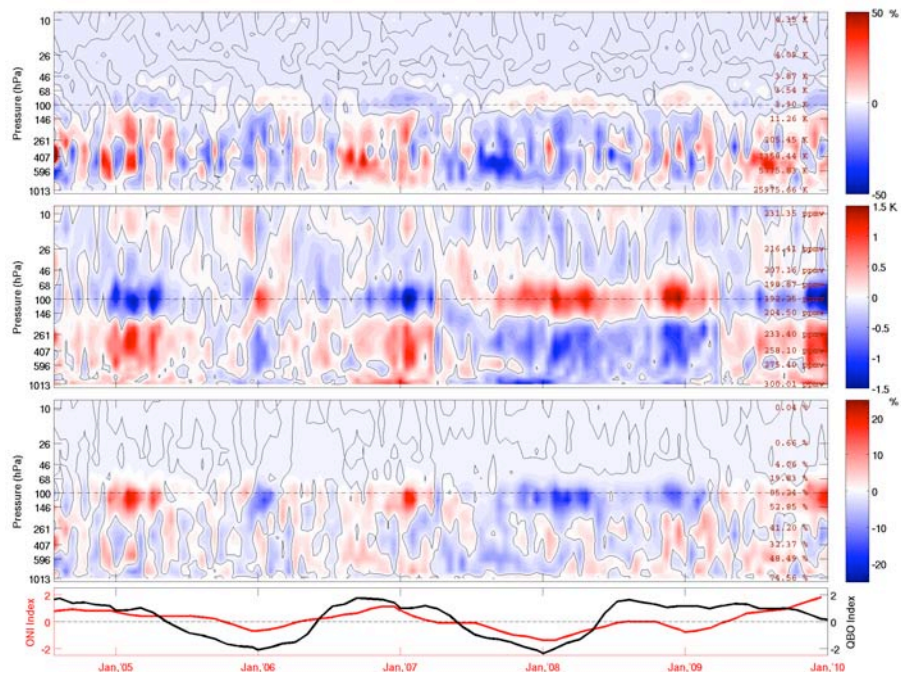

Figure 3. Same as Figure 3, but for the differences between the time series over the Niño3.4 region and the entire tropics.
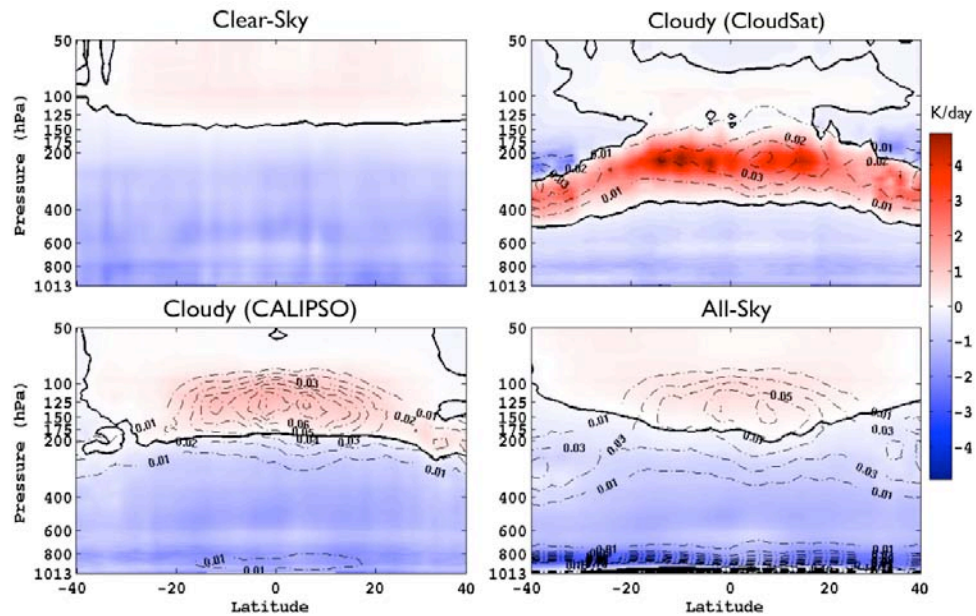
Figure 4. Longwave radiative heating rates (K/day) for (a) clear-sky, (b) scenes with only CloudSat clouds, (c) scenes with only CALIPSO clouds, and (d) all-sky for the zonal mean from 40S-40N. Note all-sky here is defined as any scene with full soundings of H_2O, temperature, and ozone (from AIRS) and no mix phased clouds. Dash-dotted contours are the frequency of cloud occurrence of the clouds that went into the RTM calculation for each of the panels. The solid contours represent the "zero" line.

CUBESAT FACILITY DEVELOPMENT FOR MULTIDISCIPLINARY SPACE SYSTEMS ENGINEERING EDUCATION AND DEVELOPMENT OF SMALL-SATELLITE PIPELINE

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1319

Richard Hofer (PI), Propulsion and Materials Engineering Section (353)
Alec Gallimore (Co-PI), Department of Aerospace Engineering, University of Michigan
James Cutler, Department of Aerospace Engineering, University of Michigan
Brian Gilchrist, Department of Electrical Engineering, University of Michigan
Peter Washabaugh, Department of Aerospace Engineering, University of Michigan
Kiril Dontchev, Department of Aerospace Engineering, University of Michigan
Kenneth Gmerek, Department of Aerospace Engineering, University of Michigan
Michael Heywood, Department of Aerospace Engineering, University of Michigan

A.   OBJECTIVES

The objectives of this educational initiative were to: 1) facilitate the development of dedicated environmental testing facilities and a dedicated flight hardware area for fabrication and integration of a CubeSat Flight Unit, 2) provide the infrastructure required for the design-built-test-fly (DBTF) courses being implemented as part of the Multi-Disciplinary Design (MD) Minor Program at the University of Michigan (UM), and 3) train and educate students by working on space systems and space subsystem level projects that will better prepare them for entry into the aerospace workforce.

B.   APPROACH AND RESULTS

1.   Thermal Vacuum Chamber Test Facility

A thermal vacuum chamber used to test spacecraft components and complete systems in a near-space environment has been brought to operational status under this educational initiative and is now being used to support flight projects at the University of Michigan. The environment within the thermal vacuum chamber reaches pressures as low as 5.0×10^{-5} torr and holds steady temperatures ranging from –40 °C to 85 °C. The test facility is designed to operate solely on building power and reusable resources (i.e., consumables such as liquid nitrogen are not used). Components of the thermal vacuum chamber are broken down into three subsystems that are each capable of working simultaneously or independently.

The pressure subsystem consists of a dry-scroll pump, a turbo-pump, a pressure gauge, and the vacuum chamber. The inside of the vacuum chamber can accommodate a 10 cm ×

10 cm $\times$ 30 cm 3-Unit (3U) CubeSat and takes an hour and a half to pump down to a base pressure of 5.0×10^{-5} torr.

The heating subsystem consists of two PID controllers, two power controllers, two 500-W infrared heaters, and two type-K thermocouples. When powered on, the PID controllers work with the power controllers to keep the thermocouples in the vacuum chamber at a temperature input by an operator. When a thermocouple detects temperatures below the desired set point, current is supplied to the heaters. The thermal vacuum heating system is capable of heating the vacuum chamber to 85 °C in less than 5 minutes and can hold the vacuum chamber at a constant temperature for as long as desired.

The cooling subsystem is a circulating coolant system that works to remove heat from the vacuum chamber. Our design differs from most industrial vacuum chambers that use consumables such as liquid nitrogen or liquid helium, which can be costly. An immersion chiller cools ethylene glycol in a reservoir outside of the vacuum chamber. A liquid demand pump then circulates the coolant into the chamber and back to the reservoir. The coolant flows through a copper plate while in the chamber. The copper plate serves as a heatsink and is the test stand for components. To improve the thermal vacuum chamber's cooling capacity, a peltier cooler and multilayered insulation will be installed onto the chamber in the near future.

A Labview data acquisition system is currently under development that will control and record temperature data within the chamber. When complete, the data acquisition system will be able to simulate thermal profiles equivalent to those a CubeSat will see in orbit.

2.   Clean Room Flight Hardware Facility

The development of a cleanroom area for flight hardware development has been another important step in developing the capabilities at the University of Michigan. For spaceflight projects like those underway at the University, students must have access to facilities that are clean enough for procedures such as soldering components to flight unit printed circuit boards (PCBs) and attaching solar cells to the flight systems. This dedicated area has been used in the past year for both of the CubeSat projects under development, the Michigan Multipurpose Minisatellite (M-Cubed) and the Radio Aurora Explorer (RAX). Many components of RAX's engineering design unit (EDU) and flight unit have been or are being developed in the cleanroom. M-Cubed will also be making heavy use of the flight hardware area in the coming months as it draws closer to integration.

3.   Multidisciplinary Design Minor Benefits and Student Career Preparation

The development of facilities at UM is a steppingstone to a larger goal of providing the necessary facilities for students in the MD Minor program to gain experience with full DBTF systems. This experience comes in two ways: 1) by developing the systems themselves that need to be tested with such facilities; and 2) to gain familiarity with the equipment and recorded measurements associated with these facilities. Gaining experience in this manner also contributes towards the education of other students in systems design. Students in both the Master's of Engineering in Space Engineering program and the Student Space

Systems Fabrication Lab (S3FL) are using these developed facilities for classroom projects and extracurricular DBTF systems.

All of these facilities also help to prepare the students who use them to enter the aerospace industry after obtaining their degrees. The experience gained by working on real flight projects with industry quality testing equipment gives students knowledge that cannot be gained from the classroom. This experience will lead to engineers graduating from the University of Michigan that are more well-prepared for their first assignments, allowing them to make larger contributions early on. As examples of how students are making use of the facilities to gain this experience, this educational initiative has already made it possible for students to test the electrical power system for RAX and payload candidates for M-Cubed (i.e., cameras) in a representative thermal-vacuum environment. The skills acquired during these test campaigns will be invaluable to these students as they begin careers in the aerospace field.

C.   SIGNIFICANCE OF RESULTS

A dependable and durable thermal vacuum chamber is crucial to a successful CubeSat program at the University of Michigan. Individual components can both be tested for failure and thermally baked-out for qualification. Seeing that a component fails in the thermal-vacuum chamber allows for design changes to reduce risk of space systems and their components. The benefit of having a thermal-vacuum chamber was seen during the test of RAX's Electrical Power System. Under high vacuum, the EPS was heated up to 80 °C, causing the battery to cease functioning. Without having to bring the vacuum chamber back up to atmospheric pressure, the chamber's temperature was brought back down and then slowly re-raised to locate the cutoff functional temperature of the battery. Such a situation shows the robust nature of this system and its ability to detect and catch possible failure points in space systems that otherwise may not be identified. Also, the new cleanroom facilities available at UM will promote the design and integration of current and future space projects. These projects will continue to educate and train students on full DBTF projects for years to come.

Additionally, all of these projects were in the scope of the MD Minor program, which aims to educate young engineers on the complex issues and challenges that are associated with developing sophisticated engineering systems. The knowledge gained by these students has also extended to other students who have worked with those in the MD Minor program, and into other areas of UM such as the Master's of Space Engineering program, and S3FL. All of this experience will prepare these students for careers in aerospace engineering, and this in turn will help to improve the aerospace field with the experience these students bring.

The University will continue to improve its test facilities, providing students the opportunity to learn the science of spacecraft design. In the future, Michigan may provide other universities with a means of design risk mitigation at low cost through the use of our test facilities. Students entering the workforce from these programs are better adapted for the demands of the aerospace industry and will help push space technology development faster and cheaper.

D.   FINANCIAL STATUS

The total funding for this task was $25,000, all of which has been expended.

E.   ACKNOWLEDGEMENTS

F.   REFERENCES

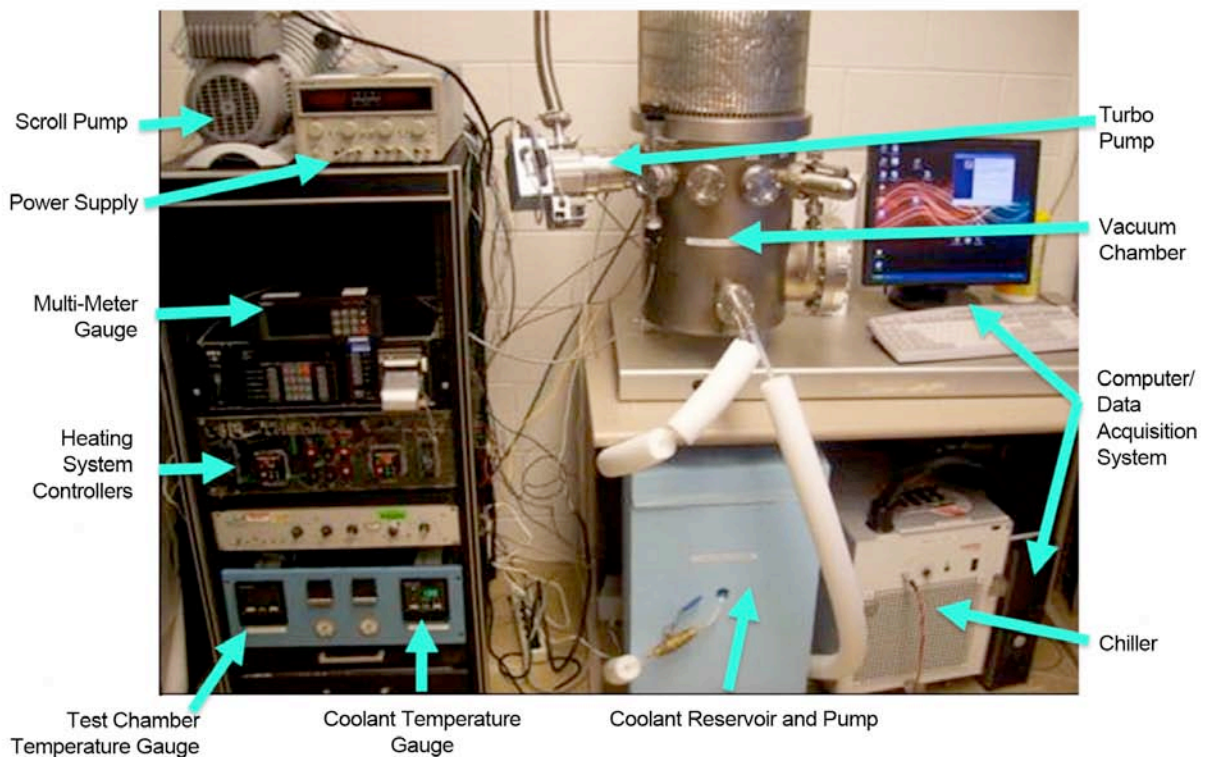None.

G.   FIGURES

Figure 1. The thermal-vacuum chamber at the University of Michigan used to support student CubeSat projects.
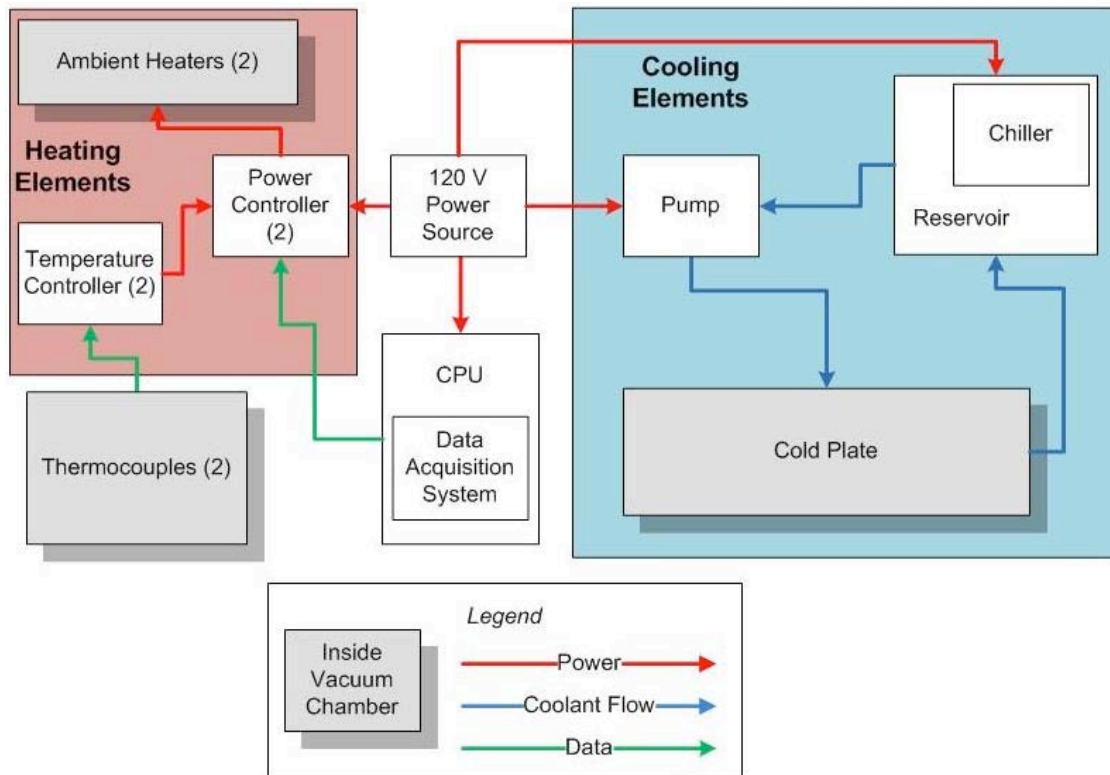
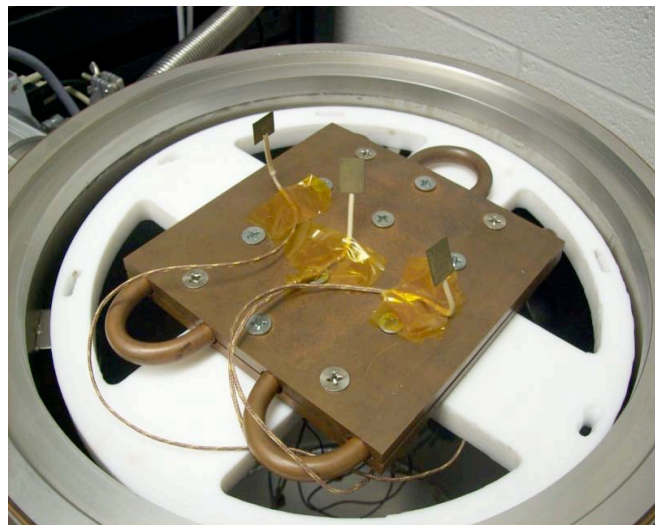**Figure 2.** System block diagram of the thermal-vacuum chamber.

Figure 3. Copper cold plate/test stand with type-K thermocouples.

Figure 4. Testing of the Radio Aurora Explorer (RAX) CubeSat in the thermal-vacuum chamber.

Figure 5. M-Cubed student in the new cleanroom facilities at Michigan.

LUNAR PROTOTYPE LANDER DESIGN TEAM: GENERATION 1

Director's Research and Development Fund (DRDF)
Final Report

JPL Task #1337

Paula J. Grunthaner (PI), University Research Affairs Office (121)
David A. Barnhart (Co-PI), Department of Astronautical Engineering and
Information Sciences Institute, University of Southern California
Joseph A. Kunc (Co-PI), Department of Astronautical Engineering,
University of Southern California

A.   OBJECTIVES

The general objectives were to educate students on a platform development team through a rigorous yet abbreviated life cycle of a systems engineering spacecraft development project, including delivering PDR/CDR and FRR review packages to outside industry representatives, culminating in an actual free flight of a simulated lunar lander.

The objective was to be enabled by building upon the FY'07/'08 USC LEAPFROG project, in which students — both undergraduate and graduate students — designed and built a Lunar Lander Prototype Flight Vehicle in a single academic year (Figure 1).

B.   APPROACH AND RESULTS

The specific plan to execute this effort was to follow an abbreviated systems engineering lifecycle of design, review, build and integrate, review, test, review and flight. Tasks and milestones for this were set up from August 2009 through May 2010 in milestones and activities in which the students could participate.

While the original plan was to concentrate on a single project lifecycle (i.e., the LEAPFROG platform), an opportunity in the 2nd semester (spring 2010) came along to design, build, integrate, and deliver a 1U CubeSat from a proprietary customer. Thus, the team was split between the development of a next-generation engine on the LEAPFROG vehicle and the 1U CubeSat design to build activity. Both are described below.

1.   LEAPFROG Activities Fall/Spring 2009

There were multiple activities that occurred during the fall and spring semesters on the LEAPFROG project. Most were centered on detailed design and implementation of the attitude control system and jet engine subsystem for the Generation-1 platform. (Gen-0 used cold gas with four thrusters and a 50-lbf engine; Gen-1 was going to use 12 thrusters with a 100-lbf engine). The detailed activities for each are shown below.

a.   Fall Semester, 2009

i.   Design and creation of a detailed test stand for the jet engine, along with hardware list (Figure 2). Prototype test stand was made and fitted with the engine. Parts were purchased and assembled for dry run of the assembly.

ii.   Worked with jet engine manufacturer on test procedures to start the engine and operate it. Significant details were missing in the original documentation, requiring us to create a complete set of documents. Also, specific parts purchased didn't work immediately due to inconsistencies in the specifications, thus some iteration with vendors had to be done over time.

iii.   Did detailed design update on the attitude control system, and tank sizing analysis. Issue is mass versus total volume of gas able to run the ACS system. Also configuration of the ACS system with multiple thrusters, and the availability of affordable valves or valve/thruster combination.

b.   Spring Semester, 2010

i.   Worked details on the procedures for the operation of the engine and testing. Systems were redesigned once again with more understanding from another source that has worked with this engine on best (and safest) way to operate the engine (i.e., a ground air source was needed to spool the engine up, and recommendation on a very large fuel reservoir for operating the engine during all phases of startup, flight thrust, and cool down).

2.   New Opportunity, 1U Cubesat Build for Spring 2010 Semester

A unique opportunity was offered to the SERC and Astronautical Engineering Department project course students to participate in a full 3U CubeSat flight development activity, by building, testing, and delivering a 1U section entirely by the end of the semester (Figures 2–6). This opportunity presented itself on 1 February, and the entire USC team delivered the functional 1U on 24 May. The 1U bus, named "CAERUS" for the Greek word for "opportunity," was to function as the 3U's primary communication subsystem.

The CAERUS bus development activity, spring 2010, required that one-half of the project team split off from LEAPFROG to help deliver the 1U in the semester. Detailed activities are described below:

i.   Develop a complete ground architecture for beacon operations. This included having to recondition an existing Yagi antenna on the roof of the physics building on campus, purchase a new computer and ground receiver, recalibrate the antenna controller, load in antenna control software based on orbital elements, and validate that everything runs with existing satellites on orbit.

ii.   Develop a "payload banger." This was developed out of LEAPFROG components as a microcontroller to provide for a payload simulator for data generation. The payload provider had not provided one, so the students developed their own.

iii.   Develop a complete computational algorithm to convert Doppler measurements per pass into range and range rate to update the altitude and orbital period of the satellite.

iv.   Communicate with worldwide beacon ground stations at various universities to coordinate the satellites beacon information being passed over the amateur radio network.

v.   Develop a brand-new quad-monopole antenna that is deployable, and validate it inside USC's Ultralab. The EE department controls the Ultralab, and we received Department approval to use it with one of our grad student teams. The antenna was tested completely in the RF chamber and characterized on an actual CubeSat structure.

C.   SIGNIFICANCE OF RESULTS

For both efforts this past year, all the students gained a significant level of experience in the effort.

Under the LEAPFROG effort, a complete set of procedures was developed, along with a test stand and setup to operate an engine that is twice has complex as the existing Generation-0 vehicle. A new electronic control unit was tested and validated as well. The students on the Generation-1 effort had to exercise both mechanical and electrical engineering skill sets, both of which challenged each set of incoming students.

Under the CubeSat effort, activities within orbital mechanics were challenged by having to go back to traditional ways of calculating orbital ephemeris with just Doppler. Additionally, a complete communications link budget and antenna analysis had to be done, much of which is not taught inside the core curriculum. The hands-on experiences were significant in developing both a data simulator for an "unknown" payload and validation of a real communications telemetry stream, from the source antenna on the satellite, all the way through a ground tracking antenna into an automated computer system on the ground.

D.   NEW TECHNOLOGY

There was no new technology developed under this effort. A new piece of technology funded by internal USC funds was tested inside the Ultralab under DRDF funding.

E.   FINANCIAL STATUS

The total funding for this task was $25,000, of which $22,577 was expended.

F. PUBLICATIONS

None.

G. REFERENCES

None.

H. FIGURES

Figure 1. USC LEAPFROG vehicle configuration. USC LEAPFROG student teams have worked on this concept from fall 2006 through the current 2009/2010 academic year.
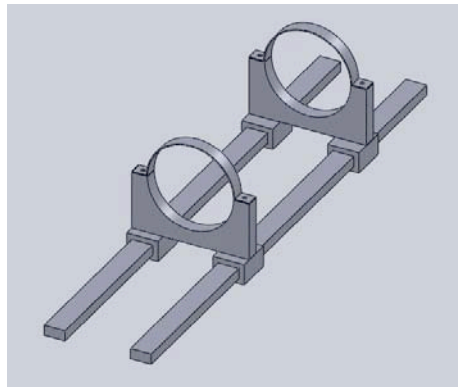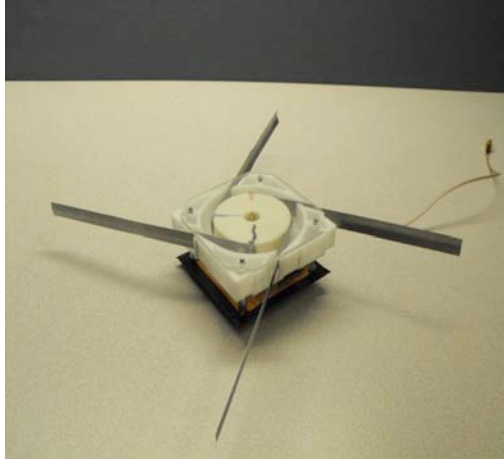

Figure 2. Generation-1 engine test stand layout.

Figure 3. New antenna tested under DRDF funding for spring 2010 CubeSat activities.

Figure 4. Existing USC antenna on physics building that was reconditioned and operating for CubeSat mission ground station.
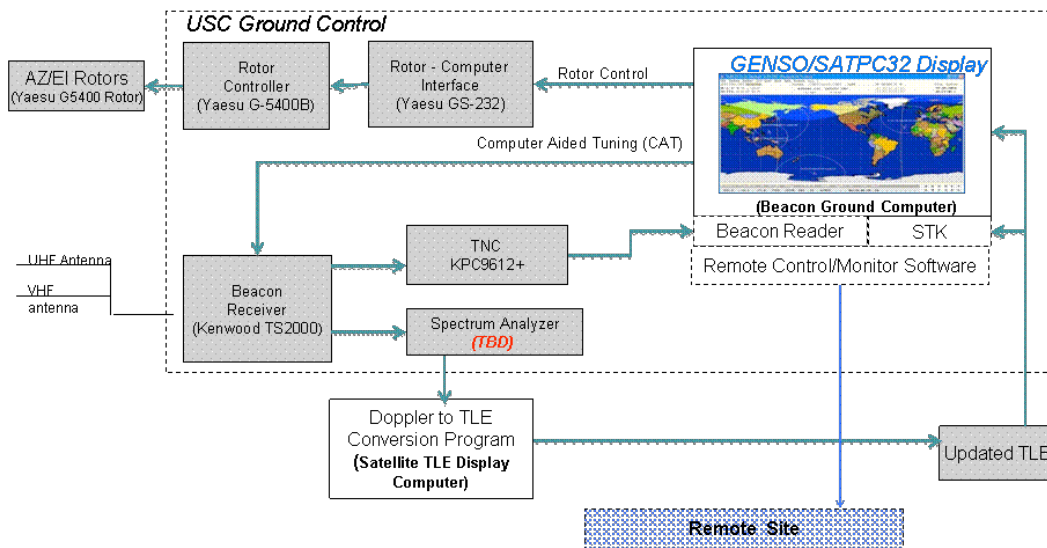
Figure 5. CubeSat mission ground station architecture developed by the students.
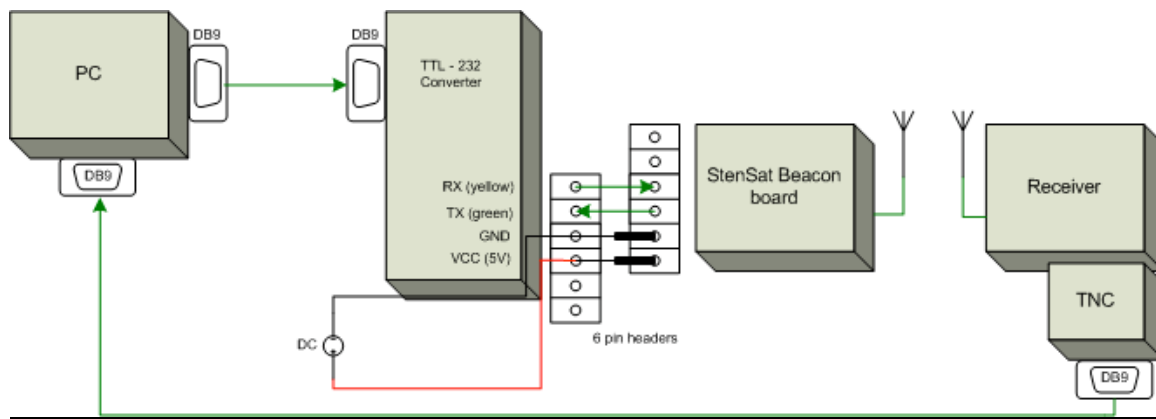
Figure 6. CubeSat mission beacon hardware in the loop test setup.

SURVIVABLE SYSTEMS FOR EXTREME ENVIRONMENTS

RADIATION-HARD, FULLY DIGITALLY CONTROLLED SINGLE-CHIP POINT-OF-LOAD REGULATOR

Philippe C. Adell (PI), Electronic Parts Engineering Office (514)
Bert Vermeire, Department of Electrical Engineering, Arizona State University
Bertan Bakkaloglu, Department Electrical Engineering, Arizona State University
K. K. Ma, Sandia National Laboratory

A.   OBJECTIVES

The objective of the work is to demonstrate the feasibility of a cost-effective, highly reliable, radiation-hardened, single-chip DC–DC switching power converter compliant to JPL missions. We designed, fabricated, and tested an innovative digitally controlled converter for point-of-load (POL) application with built-in self-test (BIST) diagnostics and load characterizations.

The goals were to:

•   Design, fabricate, and test a digitally controlled single-chip buck converter prototype using a non-hardened foundry. This collaborative research effort between JPL and Arizona State University (ASU) resulted in one of the most aggressive topologies ever reported.

•   Investigate the Sandia National Laboratory (SNL) foundry and transition the prototype to the $1B MESA fabrication facility at SNL, a strategic JPL alliance.

B.   APPROACH AND RESULTS

1.   Approach

a.   Design of the scalable DC–DC converter module

A basic representation of the converter module with direct main power connectivity, self-diagnostics, and adaptive digital control is shown in Figure 1. The digital control system operation is as follows: after voltage-to-frequency conversion using ring oscillator–based VCOs, single-bit sigma-delta (SD) modulated feedback signals are compared to the SD modulated analog voltage reference. The 3-level first-order SD noise-shaped error signal is then decimated using a 2-stage comb (CIC) filter and is applied to the compensator (PID) input. The PID compensator calculates the required duty cycle. Finally, the DPWM converts this voltage command into a duty cycle to drive the PFET and NFET via a built-in non-overlap dead-

time gate driver. The implementation of the converter architecture is based on the following modules:

• Digitally controlled DC–DC converter. A digital PWM DC–DC buck converter utilizing a first-order SD-based frequency discriminator (SDFD) is the core converter for this integrated circuit. A frequency discriminator generates an accurate representation of an instantaneous frequency of a carrier signal.

• Digital PWM generators. We use a coarse and fine scheme to generate the high-accuracy PWM duty cycle. The coarse scheme uses a counter-based approach, while the fine control is through a high-accuracy, phase-locked, digitally controlled ring oscillator, i.e., DLL. This DLL controls the 4 fine bits of DPWM codes. To minimize the effect of single-event or other harsh environment, existing critical blocks in analog DLL such as phase frequency detectors, charge pumps, and voltage-controlled oscillators or delay lines are replaced by a digitally intensive, multiplying delay-locked loop (MDLL) circuit. With this approach, digital cell libraries and hardening techniques, such as temporal redundancy, can be utilized for the clock-multiplying DLLs.

• Digital self-diagnostic circuits and optional digital current sensing techniques. As shown in Figure 2, a self-diagnostic circuit is used to measure the Resr and L values of the power inductor during converter operation. We built an on-chip triangular current generator, and pushed it into the inductor while shorting the power capacitor. An OPA, VCO, and SDFD chain is used to read out and digitize the voltage over the inductor. The readout chain is similar with the VCO SDFD chain in the digital control loop shown in Figure 2.

b. Transition to a non-hardened foundry: assessment

To match JPL mission requirements and because power converters are very sensitive to total ionizing dose and single-event effects, we studied the feasibility of transitioning the current converter prototype to the SNL rad-hard foundry. The design is currently being manufactured using the rad-hard process, which offers unique capabilities that combines CMOS transistors with high-voltage drivers. The proposed design will leverage the commercial development and will target complete total dose and single-event hardness, a digital control implementation that is fully synthesizable including the DPWM generation and load current monitoring.

2. Results

Two generation designs are developed during the course of this project. A second-generation design from FY'09 development is currently being manufactured using a high-voltage commercial foundry. The innovative features are 1) higher voltage and current requirements (i.e., 12 volt, 2 A), and 2) a digitally intensive MDLL circuit, and 3) fully synthesizable. The radiation-hardened version of design is currently being transitioned to the SNL foundry.

High-voltage design. The DPWM generation and digitally intensive feedback digitizers have been successfully demonstrated experimentally. The converter results are shown

in Figures 3 and 4. Figure 3 shows a closed-loop measurement; the converter output was recorded and regulated properly. The output ripple value was well below 1%. The converter transient output characteristics, with a step load change, recovered within 85 ms and its transient change stayed within 5% of the output value. Figure 4 shows the results of the inductor current measurements using the BIST measurement technique. With this technique, we were able to measure inductance values from 2 to 22 mH with very good accuracy (within a few percent). Figure 5 shows a layout picture of the chip; the BIST takes only 6% of the full chip area (about 12 mm^2). Figure 6 shows the efficiency plot, indicating a maximum peak efficiency of 92% at 1 A, a very competitive number knowing that the process is not optimized. Finally, we performed a single-event transient test using our new JPL laser pulsed system. The converter showed very good worst-case transient performance (see Figure 8) and the most sensitive portion of the circuit has been identified. The high-voltage version is currently being fabricated; it will incorporate the new features described above, i.e., higher voltage, MDLL to reduce transient effects, and fully synthesizable.

Sandia process evaluation. Sandia offers a unique and attractive 1 Mrad process for the current approach. It integrates 0.35-µm silicon-on-insulator (SOI) transistors with high-voltage drivers offering low specific-on-resistance N-channel LDMOS devices at up to 30 V. Intensive efforts have focused on the process design kit to obtain a user-friendly design flow. The design has been simulated and is currently being laid out. The expected fabrication submission will be January 21st. Final results will be provided in the final report.

Process characteristics and radiation data are summarized below:

LDMOS device — A very attractive lateral DMOS device with a breakdown voltage of 30 V, 10 A, and a good gate charge-on-resistance figure of merit (150 nC.mW). FY'09 development led to a device that is immune to single-event latchup (SEL), burnout (SEB), and gate rupture (SEGR); with a threshold LET (LET$_{th}$) > 80 MeV.cm^2/mg. These improvements impacted total hardness; however, as shown in Figure 9, the device shows little degradation up to 100 krad and the threshold voltage is still positive after 500 krad up to 1 Mrad.

CMOS7 — A rad-hard, mixed-signal CMOS SOI (BUSFET) technology, 0.35 mm and 3.3 V operations. As shown in Figure 10, 1) total dose up to 500 krad with no parametric degradation and >1 Mrad with parametric degradation, no functional failure; 2) single-event effects (SEE) SEU threshold LET ~ 30 MeV.cm^2/mg.

C.   SIGNIFICANCE OF RESULTS

The industrial supply base for space-grade DC–DC converters has had routine difficulty meeting the reliability, schedule, and manufacturing standards set forth by the space industry. This effort develops a digitally controlled single-chip solution, thereby eliminating the principal production problems associated with the customary hybrid construction. The targeted 300 krad (Si) (goal 1 Mrad (Si)) power converter, with its improved reliability and power efficiency, will be industry leading. Although all missions will benefit from the reduced cost and schedule risk typically associated with commercial power converters, this design will be crucial for the planned Europa mission. Its scalable digital control approach lends itself to future applications

that will require higher bus voltage (i.e., 70 V). We are targeting a design solution that can be used as an alternative to hybrid COTS converters in JPL power-distribution units with the advantages of design flexibility, with similar or even better performance and reliability, and radiation-hardened. High-efficiency POL converters will greatly improve the overall efficiency of power systems, thereby significantly reducing power requirements for all JPL missions.

E.   FINANCIAL STATUS

The total funding for this task was $100,000, all of which has been expended.

F.   ACKNOWLEDGEMENTS

G.   PUBLICATIONS

[A]   T. Liu, H. Yeom, B. Vermeire, P. Adell, and B. Bakkaloglu, "A Digitally Controlled DC–DC Buck Converter with Lossless Load Current Sensing and BIST Functionality," accepted as an oral presentation at the IEEE International Solid-State Circuits Conference, San Francisco, California, February 20–24, 2011.

H.   REFERENCES

None.
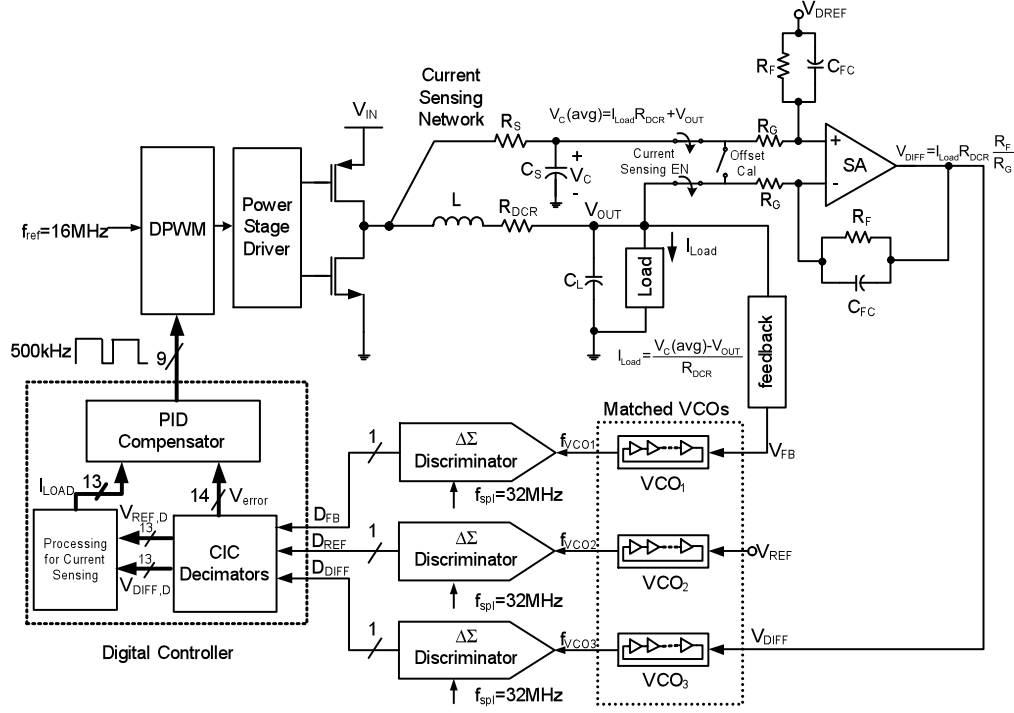
I. FIGURES

Figure 1. Proposed DC–DC buck converter architecture with lossless load current sensing circuitry.
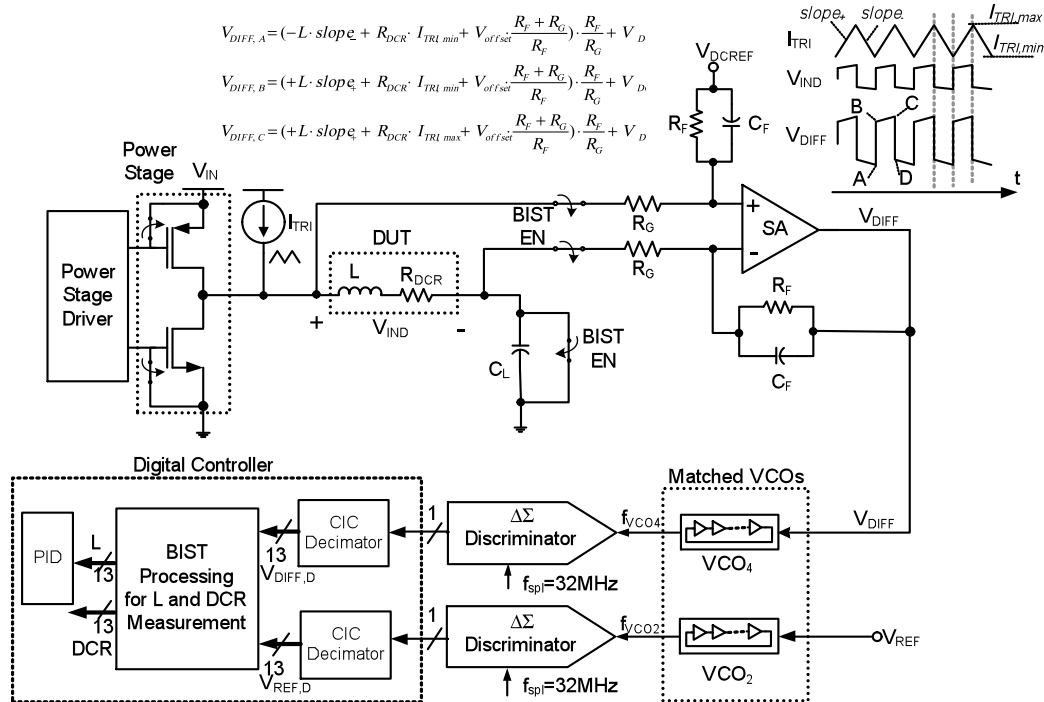

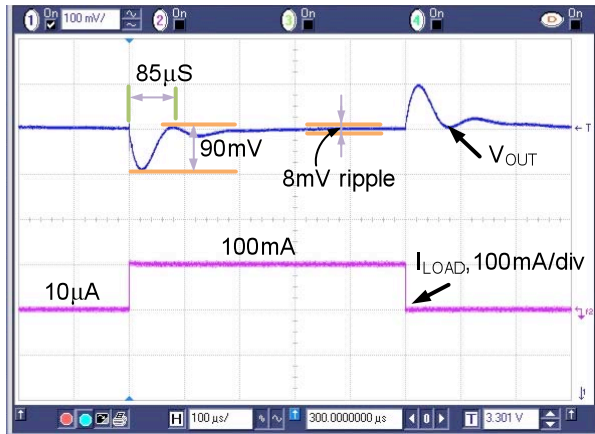**Figure 2.** BIST mode signal chain for inductor inductance and DCR characterization.

Figure 3. DC–DC buck converter transient response for 100 mA step, 1 kHz switching load current. VIN = 5 V, VOUT = 3.3 V, CL = 22 μF, ESR = 70 mΩ, L = 18 μH, DCR = 63 mΩ.
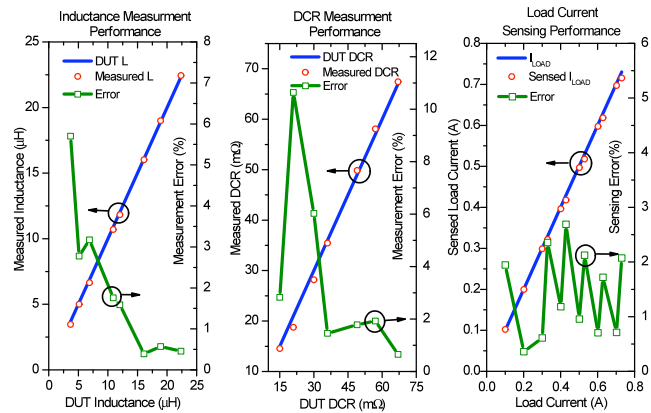
Figure 4. Inductor measurement and load current sensing performance, achieving 2.1% average error for inductance measurement, 3.6% average error for DCR measurement, and 1.5% average error for load current sensing. Here, current sensing is based a DCR = 63 mW, L = 18.8 mH inductor case.
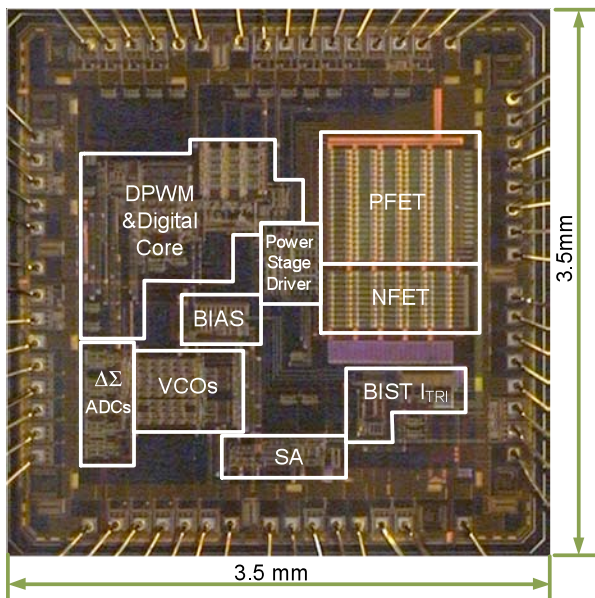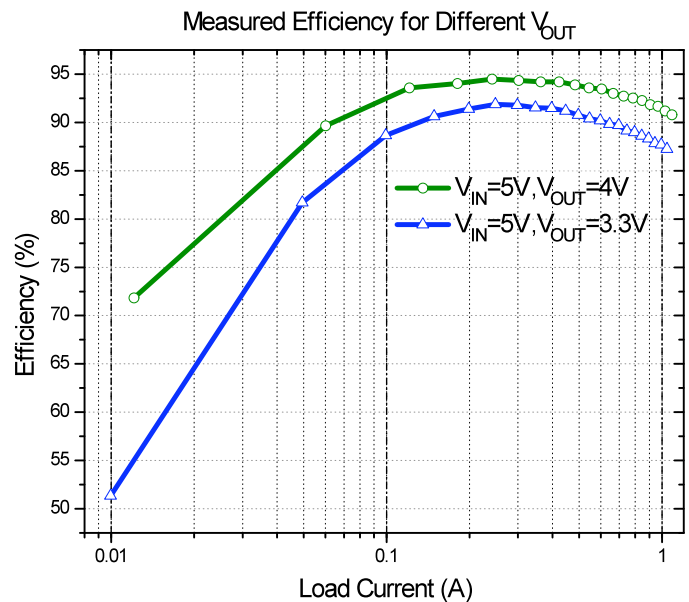

Figure 5. Die micrograph.

Figure 6. Measured DC–DC buck converter efficiency for different VOUT@ VIN = 5V.
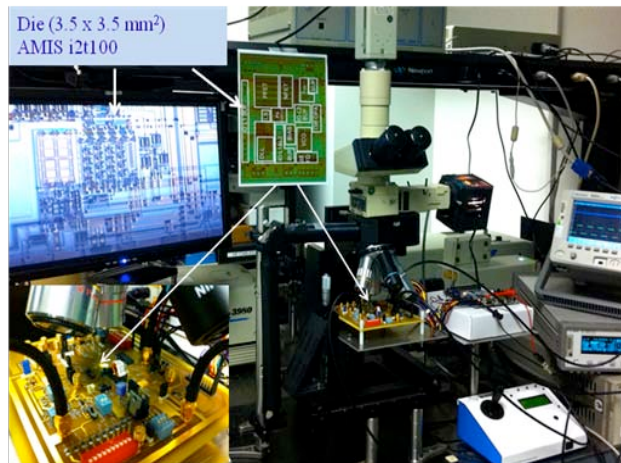
Figure 7. JPL laser experimental setup. The POL was tested using a 20-pJ laser pulse. The main sensitive region of the device was the DLL.
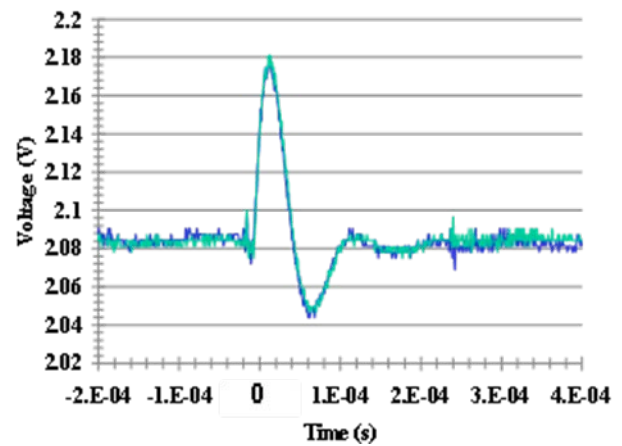

Figure 8. Typical worst-case transient measured at the ASU–JPL POL converter output. Transient is well within the load step change transient response.